UNIVERSITY
of
GLASGOW

Characterisation of the virulence-related, outer-membrane proteins of *B. pertussis*

Paul Edward Blackburn B.Sc.

Presented for the degree of Doctor of Philosophy in the Division of Infection and Immunity, Institute of Biomedical and Life Sciences, Faculty of Science, University of Glasgow

May 2000

ACKNOWLEDGEMENTS

I would like to thank my supervisors, Dr Roger Parton and Dr John Coote for their excellent supervision, proof-reading of this thesis and their faith in me. Thanks also to Professor Mark Roberts for his help with the codon usage predictions and general enouragement throughout this project. I would also like to thank Professor A.C. Wardlaw for his help with the statistical analysis of the mouse protection data. I am grateful for technical guidance from Susan Campbell. I would like to thank the division as a whole for being supportive and allowing my study to be enjoyable.

I appreciate the emotional and financial support of my parents, Mick and Jane, for their absolute confidence in my abilities, and for making it possible for this thesis to be undertaken. I would also like to thank my sister Helen and my nephew Cameron for providing entertainment during visits to Harrogate.

Many thanks to Kim Duncan for her untiring support and enthusiasm during my study, in particular her good humour when things were tough.

Finally, I would also like to thank the many good friends I have made during the course of this study, especially those that have supported me through difficult times and made the good times even better.

CONTENTS

List of figures

List of Tables

ABBREVIATIONS

ACV	Acellular vaccine
AL	Alhydrogel
Anova	Analysis of variance
ATP	Adenosine triphosphate
BLAST	Basic local alignment
BG	Bordet-Genogou
BrkA	*Bordetella* resistance to killing A
BSA	Bovine serum albumin
Bvg	*Bordetella* virulence gene
CAA	Casamino acids
CL	Cyclodextrin liquid
Da	Dalton
DNA	Deoxyribonucleic acid
FHA	Filamentous haemagglutinin
FPLC	Fast protein liquid chromatography
GSP	General secretory pathway
h	Hour
HLT	Heat-labile toxin
IPTG	Isopropyl ß-D-thiogalactopyranoside
KDa	KiloDalton
Kpsi	Kilo-pounds per square inch
KV	KiloVolts
l	Litre
LB	Luria Bertani
LOS	Lipooligosaccharide
LPS	Lipopolysaccharide
M	Molar
MM	Minimal media
mM	Millimolar
min	Minute

ml	Millilitres
µl	Microlitres
Mw	Molecular weight
Ni-NTA	Nickel-nitrilotriacetic acid
nl	Nanolitre
nm	Nanomolar
°C	Degrees Celsius
OMP	Outer-membrane protein
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
POMPS	Processed outer-membrane proteins
Prn	Pertactin
PT	Pertussis toxin
P.30	30 KDa processed form of pertactin
P.69	69 KDa processed form of pertactin
P.93	93 KDa unprocessed form of pertactin
RGD	Arg-Gly-Asp
RNA	Ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
s	Seconds
SDS	Sodium dodecyl sulphate
Sec	Secretory
Tcf	Tracheal colonisation factor
TCT	Tracheal cytotoxin
Vag	Virulence-activated gene
Vag8	Virulence-activated gene 8
Vrg	Virulence-repressed gene
WCV	Whole-cell vaccine
x *g*	Gravity
X-GAL	5-bromo-4 chloro-3-indolyl-ß-D-galactodisase

ABSTRACT

At the beginning of this study, knowledge of autotransporter proteins was in its infancy and the *B. pertussis* proteins pertactin , BrkA and tracheal colonisation factor (Tcf) had not been recognised as members of this family. However, the sequence similarity between these *Bordetella* proteins and a number of other proteins from different genera soon became apparent and the *Bordetella* proteins were also included in the autotransporter protein family. It was the similarity between pertactin and other *B. pertussis* proteins that, due to a mis-primed PCR amplification, led to the preliminary data for this project. A *B. pertussis* genomic library was screened, using this amplicon as a Southern blot probe, and a cosmid was identified and subcloned. Analysis of the sequence led to the identification of a larger open reading frame which had not been previously recognised. The cosmid did not contain the entire open reading frame and a further 200 bp was obtained using an adapter-linked single specific primer-PCR (SSP-PCR). An open reading frame was predicted and the amino acid similarity strongly suggests that this gene product represents another member of the autotransporter family. This protein was named Bap-5 as the fifth member of the *Bordetella* autotransporter protein family. Motif analysis revealed an RGD integrin-binding motif and an SGSG glycosaminoglycan-binding motif, which implicates this protein as a putative adhesin. No upstream regulatory sequences were identified although a Rho-independent terminator sequence was predicted downstream. The *bap-5* sequence was deposited in GenBank on 11/08/1998 (accession number AF081494).

Although not begun at the start of this study, the *B. pertussis* genome sequence was near completion during the final stages of this work and a sequence identical to *bap-5* was located in the sequence data. A similar sequence to *bap-5* was also identified in the genome of *B. bronchiseptica*, for which a sequencing project was initiated during this study. Progress was made towards the construction of a *bap-5* deficient mutant, although the final stages were not completed. Such a mutant would allow the role of

Bap-5 to be characterised. Sequences which hybridised to a *bap-5* Southern blot probe were identified in *B. bronchiseptica* and *B. parapertussis* but not in *B. avium*. The expression of Bap-5 was confirmed by RT-PCR and Western blotting, and shown to be Bvg-regulated. Such regulation implies a role for Bap-5 in pathogenicity of *B. pertussis*. The Westen blotting results also suggested that Bap-5 is processed in a similar way to pertactin.

Recombinant expression systems were set up for Bap-5, a specific portion of Bap-5 and the carboxy-terminal domains of pertactin, BrkA, Tcf and Bap-5. Most expression vectors used allowed the construction of histidine fusion proteins which could be readily purified. These autotransporter domains provided material for mouse protection tests and antibody production. Under the conditions tested, however, these recombinant strains did not protect against lung colonisation after intranasal challenge with *B. pertussis* although mature pertactin (P69) gave some protection. In addition, a heterologous antigen display system, based on the pertactin sequence was created. A sec-dependent signal sequence was fused to the pertactin C-terminus along with some upstream sequence which allowed targeting of this domain to the outer membrane of *E. coli*. Future development of this vector, *ie.* fusion of a passenger domain which could be displayed on the bacterial surface, may prove useful in vaccine or other biotechnological applications.

1.0 Introduction

1.1 Genus and species characteristics

1.1.1 Definition

Bordetella species are small, aerobic, non-sporing Gram-negative cocco-bacilli (0.3-0.5 μm x 0.5-2.0 μm). The optimum growth temperature for bordetellae is 35-37°C and carbohydrates are not fermented (Parton 1998). The genus *Bordetella* comprises a closely-related group of pathogens that are capable of infecting the respiratory tracts of warm-blooded vertebrates. With the exception of laboratory cultivation, they are considered to be obligate respiratory tract pathogens. However, the survival and growth of *B. bronchiseptica* and *B. parapertussis* in natural waters and other low nutrient conditions suggests that these species have an environmental reservoir (Porter *et al.* 1991; Porter and Wardlaw 1993). The primary habitat of *B. pertussis* is the surface of the ciliated epithelial cells in the human respiratory tract although it has been suggested that invasion of mammalian cells may provide an alternative, intracellular niche (Lee *et al.* 1990; Friedman *et al.* 1992; Forde *et al.* 1998).

1.1.2 Taxonomy

Bordetella species have been previously assigned to various different bacterial genera including *Alcaligenes*, *Bacillus*, *Haemophilus* and *Brucella* (Parton 1998). A new family, *Alcaligenaceae*, was suggested on the basis of DNA-rRNA hybridisations and phenotypic analyses (De Ley *et al.* 1986). Recent evidence from phylogenetic 16S rRNA sequence analysis has confirmed that *Bordetella* most closely resembles *Alcaligenes* and these genera are now assigned to the beta-2 subdivision of *Proteobacteria* (Weyant *et al.* 1995). To date, seven species have been included in the genus *Bordetella*. *B. pertussis* (originally *Bacillus pertussis*) was identified as the cause of whooping cough (pertussis) in 1906 (Bordet and Gengou 1906). *B. bronchiseptica* (originally *Bacillus bronchicanis*) has been recognised as a cause of kennel cough in dogs

and also causes respiratory disease in other wild and domestic mammals (Ferry 1910). Occasionally, *B. bronchiseptica* is responsible for causing disease in man. *B. parapertussis* (originally *Bacillus parapertussis*) is responsible for causing a generally mild form of whooping cough in man (Eldering and Kendrick 1938) and related strains have been isolated from the respiratory tracts of sheep (Cullinane *et al.* 1987; Khelef *et al.* 1993). *B. avium* was isolated as a *B. bronchiseptica*-like organism and the causative agent of turkey rhinitis (Filion *et al.* 1967) although this species designation was not suggested until 1984 (Kersters *et al.* 1984).

Recently, three more species have been included in the genus *Bordetella* on the basis of genotypic and phenotypic comparisons. *B. holmesii* was isolated originally from the blood cultures of patients, some of whom were immunocompromised (Weyant *et al.* 1995). *B. hinzii* was the name proposed for a non-pathogenic *B. avium*-like group of organisms that has been isolated from the respiratory tract of chickens and turkeys and also from immunocompromised humans (Cookson *et al.* 1994; Vandamme *et al.* 1995). The most recent addition to the genus is *Bordetella trematum*, which has been isolated from wounds and ear infections in humans (Vandamme *et al.* 1996). There is little information regarding the properties and characteristics of these three recently-described species.

1.2 Clinical disease-pertussis

Pertussis is a highly communicable disease and has an attack rate of over 90% among unimmunized individuals (Lambert 1965). Although *B. pertussis* is a world-wide pathogen, an intensive vaccination programme has proved effective in drastically reducing the number of pertussis cases in developed countries. It has been estimated that the vast majority of unvaccinated children are infected with *B. pertussis* by the age of five years (Fine and Clarkson 1987). The disease may be life threatening in young infants under 6 months of age and the variable severity of disease is influenced by the age, general health and immune status of the patient (Parton 1991).

Typical pertussis disease can be divided into three symptomatic stages, 1) catarrhal, 2) paroxysmal and 3) convalescent (Cherry *et al.* 1988). Following infection with *B. pertussis* there is an incubation period of 6 to 20 days (Lapin 1943). The catarrhal stage lasts 1-2 weeks and symptoms include a mild cough. During the paroxysmal stage (1-20 weeks) the cough is more forceful and more frequent. In approximately 50% of patients coughing paroxysms often end in a characteristic inspiratory "whoop", sometimes with vomiting, apnoea and cyanotic attacks which are the most frequent cause of pertussis hospital admissions (Cherry *et al.* 1988). Most complications, for example brain damage which is thought to be caused by cerebral anoxia, and also deaths occur at this stage. The convalescent stage is characterised by a reduction in the number and frequency of paroxysms, which may still occur sporadically for up to 6 months post-infection.

1.3 Mechanisms of pathogenesis

Bordetella spp. produce a plethora of pathogenic factors that are co-ordinately regulated during infection. A number of toxins, such as pertussis toxin and tracheal cytotoxin are secreted from the *B. pertussis* cell surface and these proteins have unusual activities associated with them. In addition *B. pertussis* also produce a range of adhesins, some of which may also act as invasins.

Proteins suggested to be involved in adhesion include fimbriae (Robinson *et al* 1990), filamentous haemagglutinin (FHA) (Stibitz *et al.* 1988), tracheal colonisation factor (Tcf) (Finn and Stevens 1995), Bordetella resistance to killing protein A (BrkA) (Fernandez and Weiss 1994) and pertactin (Prn) (Charles *et al.* 1989). The list also includes pertussis toxin which, although secreted from the cell, still facilitates adhesion and presumably must, therefore, also be surface located (Sandros and Tuomanen 1993). Flagella on the surface of *B. bronchiseptica* also appear to play a role in adhesion to eukaryotic cells (Leigh *et al.* 1993; Savelkoul *et al.* 1996). It is likely that other yet uncharacterised proteins facilitate adhesion and therefore colonisation of the respiratory

mucosal surface. Adherence appears to be multifactorial and relies on a number of outer-membrane components *i.e.* those anchored to the membrane, and cell-associated proteins (Leininger *et al.* 1990). It is not known why *B. pertussis* produces so many different adhesins, although it is tempting to speculate that they are expressed at different stages in the pathogenesis of *B. pertussis* infection and therefore allow adaption to adhere to, or release from, various surfaces during the infection cycle. Possession of an array of adhesins may enable the bacteria to adhere to host cells when antibodies to specific adhesins are present. This is known as "immune exclusion" (Weiss 1997).

1.3.1 Regulation of virulence

The expression profile of bordetellae can change dramatically *in vitro* in response to growth conditions. This adaption may be indicative of their potential to colonise, and potentially invade, distinct areas of the respiratory tract, to evade immune responses (by altering the antigens expressed) and possibly to inhabit other unrecognised niches (Parton 1991; Cotter and Miller 1996). Co-ordinate regulation of proteins is presumed to be important in the life cycle of *B. pertussis* but at what stage is unclear.

Bordetella species are able to switch between virulent (Bvg^+) and avirulent (Bvg^-) forms. The *bvgAS* locus (*Bordetella* virulence gene) encodes a two component regulatory system which is responsible for the control of many virulence-activated genes (*vags*), some of which encode the virulence factors outlined in Table 1. Several virulence-repressed genes (*vrgs*) have also been identified but are not well characterised (Stibitz *et al.* 1988; 1989; Coote 1991). The BvgS protein is located in the cytoplasmic membrane and acts as a sensor of environmental signals.

Table 1. Virulence factors of bordetellae.

Virulence Factor	*Bordetella pertussis*	*Bordetella parapertussis*	*Bordetella bronchiseptica*	*Bordetella avium*	Suggested Role in Pathogenicity
Tracheal cytotoxin	+	+	+	+	Ciliostasis, epithelial cell cytotoxicity
Heat-labile toxin	+	+	+	+	Local inflammatory effects
Endotoxin	+LOS	+LPS	+LPS	+NK	Pyrogenicity
Fimbriae	+	+	+	+	Adhesin/invasin
Filamentous haemagglutinin	+	+	+	-	Adhesin/invasin
Adenylate cyclase toxin	+	+	+	-	Interference with immune effector cells
Pertactin	+	+	+	NK	Adhesin/invasin
BrkA	+	- *	+	-	Adhesin/invasin, serum resistance, resistance to antimicrobial peptides
Tracheal colonisation factor	+	-	-	NK	Adhesin
Pertussis toxin	+	-	-	-	Adhesin/invasin, interference with immune effector cells
Osteotoxin	-	-	-	+	Cytotoxicity

NK, not known; LPS, lipopolysaccharide; LOS, lipooligosaccharide. (Parton 1996)

* Homologous sequence present

These signals are then transferred to the cytoplasmic BvgA protein by phosphorylation. BvgA then activates the transcription of the *vags* and represses the expression of the *vrgs*. The BvgA binding site has been identified upstream from the open reading frames of several virulence-associated genes including adenylate cyclase, pertussis toxin, FHA and pertactin (Karimova, *et al.* 1996; Karimova and Ullmann 1997; Kinnear *et al.* 1999). This global control of virulence factors appears to be autoregulated (Melton and Weiss 1989; Scarlato *et al.* 1990; Scarlato *et al.* 1991).

The expression of several *Bordetella* virulence factors is dependent on the phase and mode of the organism, which thus reflects the virulence of the strains. The two mechanisms involved have been termed phase variation and antigenic modulation. Phase variation results in the production of an avirulent strain which is due to a frameshift mutation or a deletion in the *bvg*S gene and occurs at a frequency of 1 per 10^3-10^6 organisms. Avirulent strains that result are described as Phase IV, whereas virulent strains are described as Phase I. This process is rarely reversible. Antigenic modulation, however, is readily reversible and also results in a change in virulence state (Lacey 1960). In antigenic modulation an avirulent form (cyanic or C-mode) results from a virulent form (xanthic or X-mode) as an *in vitro* response to growth conditions such as high levels of certain salts, organic acids or at low temperatures. The BvgAS regulatory system is thought to be responsible for the alteration of expression which is observed antigenic modulation. Phase I and X-mode strains strains are phenotypically identical, as are Phase IV and C-mode strains.

1.4 Vaccination against pertussis

Immunisation of infants with inactivated whole-cell vaccines (WCVs) has, in countries with good vaccination coverage, dramatically reduced the morbidity and mortality from pertussis. However, in recent years there has been an increase in the incidence of pertussis, particularly in older children and adults, in a number of countries including the

US and several countries in Europe. The reasons for this increase are unknown, although immunity as a result of childhood vaccination decreases after about 12 years of age. According to World Health Organisation (WHO) statistics, the number of deaths due to pertussis worldwide has been estimated at 450,000 for the year 1997 and a further 700,000 were prevented by vaccination (further details can be found at the WHO website http:www.who.int/gpv-surv/graphics/Jan99graphics/estdeaths.htm). Although whole cell vaccination programmes against pertussis have been very effective in reducing the incidence of this disease, there are concerns that current WCVs do not prevent infection (Willems and Mooi 1996). The circulation of *B. pertussis* in the human population has not been significantly affected by the WCVs (Fine and Clarkson 1987; Grenfell and Anderson 1989). This may lead to those with waning immunity becoming a significant reservoir of infection for others who are susceptible (Aoyama *et al.* 1992). Such a reservoir may explain the peaks in pertussis every 3-4 years which occur in countries with high vaccine uptake (Fine 1988).

Frequent, but usually mild, adverse reactions, and a fear of rare but serious neurological events associated with WCVs, and an increase in the understanding of the pathogenesis of pertussis have prompted the development of a new generation of better defined pertussis vaccines, the acellular vaccines (ACVs). Acellular vaccines have been used in Japan since 1981 and have been shown to be effective and have fewer side effects compared to WCVs (Kimura and Kuno-Saki 1990). Early double-blind, placebo-controlled ACV trials in Sweden (1986-1987) were performed with either monocomponent chemically-inactivated pertussis toxoid (PTd) or a two component PTd and filamentous haemagglutinin (FHA) vaccine (Ad Hoc Group 1988). The reactogenicity of these vaccines was low but the efficacy was less than would have been expected for WCVs. More recent vaccine trials have highlighted the importance of including other components (Plotkins and Cadoz 1997). Although the composition of current ACV formulations varies widely between manufacturers, important components include PTd, FHA, P.69 and fimbriae (fim 2 and 3) (Willems *et al.* 1998).

The results of recent clinical trials indicate that ACVs offer the prospect of reduced reactogenicity relative to WCVs and for some preparations may achieve comparable efficacy (Cherry 1997; Halperin 1999). One drawback of current acellular vaccine preparations is that, like WCVs, infection with *B. pertussis* does not appear to be prevented and therefore the problem of a reservoir in the human population remains.

Acellular vaccines have been shown to predilict a Th2 response, whereas intracellular infection of macrophages with *B. pertussis* or stimulation with killed *B. pertussis* induces an IL-12 driven Th1 response which is more protective in such murine models (Mahon *et al.* 1996). Therefore the use of IL-12 as an adjuvant is under study in ACV trials (Mahon *et al.* 1996).

It is evident from a recent Dutch trial that the optimal composition of ACV preparations has not yet been found (Mooi *et al.* 1998). The whole-cell vaccine used in this Dutch vaccination programme appears to have driven antigenic shifts in the PT and P.69 antigens within the *B. pertussis* population. This immune selection has caused the re-emergence of pertussis in highly vaccinated populations and may have major implications for the long term efficacy of both whole cell vaccines and acellular vaccines. These findings highlight the limitations of clinical trials which have already led to the use of several ACV preparations. It is essential that further virulence factors are sought as acellular vaccine components in an attempt to combat such problems.

1.5 The nature of the bacterial outer membrane

Most of the important virulence factors and vaccine components of *B. pertussis* are either surface located or exported (Table 1). It has been demonstrated that the envelope protein profile of bordetellae alters markedly following exposure to different growth conditions and this is related to the alteration of virulence state (section 1.3.1) (Wardlaw *et al.* 1976; Ezzell *et al.* 1981). Therefore the production, assembly and export of

surface components of *B. pertussis* are important subjects for study. Many features of the *B. pertussis* outer membrane are common to Gram-negative bacteria in general.

The distinction between outer-membrane proteins and secreted proteins is not clear-cut. Secreted proteins such as pertussis toxin (PT), FHA and Tcf have been shown to function as adhesins and must therefore be associated with the bacterial outer membrane in some way. Other surface-located proteins, such as pertactin, are perceived as outer membrane proteins, although proteolytic processing probably allows release from the C-terminal anchor domain.

The outer membrane of Gram-negative bacteria provides a structural and functional interface between the environment and the interior of the cell. Initially the outer membrane was perceived as a simple and static lipopolysaccharide-containing layer, although intense study over the last thirty years has shown the outer membrane to be a dynamic entity that is intimately involved with many cellular processes (Hancock 1991). The asymmetrical character of the outer membrane, with the outer monolayer containing heterogeneous LPS molecules, provides a unique barrier and a site of association with proteins (Moran 1996). The outer membrane functions as an exclusion barrier with selective permeability which allows the control of substances moving into and out of the cell. In Gram-negative bacteria, the major porin is involved in such control (Hannah *et al.* 1990). This porin consists of complexes of a monomer of approximately 40 KDa that assembles in the outer membrane and allows the passage of ionic molecules (Armstrong and Parker 1986; Li, *et al.* 1991). The *B. pertussis* major porin has been characterised (Hannah *et al.* 1990). The porin prevents passage of hydrophobic and large hydrophilic molecules into the cell, and numerous proteins, peptides, carbohydrates and hydrophobic drugs are actively exported by other means from the cytoplasm into the extracellular medium (Dinh *et al.* 1994).

The presence of an outer membrane is therefore very useful, but such benefits come at a price. Mechanisms must exist for the translocation of excreted proteins and the outer

membrane must also provide anchor facilities for external structures such as adhesins and transport channels (Wandersman 1992; Pugsley 1993).

1.6 Protein translocation to the cell exterior

It is clear that no one mechanism is responsible for protein translocation, and those that have been suggested have been reviewed in detail elsewhere (Pugsley and Schwartz 1985; Hirst and Welch 1988; Lory 1992; Nikaido and Saier 1992; Wandersman 1992; Pugsley 1993; Salmond and Reeves 1993). The transport of proteins across the Gram-negative envelope has been categorised into five pathways:

1.6.1 Type I secretion

This group of sec-independent proteins mediate their own transport via several accessory proteins which form a transmembrane channel through which the secreted molecule moves. This pathway is totally independent of the sec-dependent pathway and no typical signal sequences can be identified on the protein. Type I machinery consists of an ATP-binding cassette transporter, an outer membrane component related to the TolC protein of *E. coli*, and a dimeric protein which spans the cytoplasmic and outer membranes of the cell. The prototype member of this family is the *E. coli* haemolysin (Femlee *et al*. 1985). *B. pertussis* adenylate cyclase toxin and FHA are also transported via this secretion mechanism (Coote 1992). The secretion of FHA is described in more detail in section 1.8.7. The C-terminal domain of proteins which use this pathway encodes the information neccessary for secretion.

1.6.2 Type II secretion

This is thought to be the major export pathway in most Gram-negative organisms and is known as the "general secretory pathway" or "sec-dependent pathway" as it is

dependent on several sec proteins (Sec A, B, D, E, F and Y). The amino terminal region of type II-secreted protein precursors are homologous and act as a signal sequence which plays a role in delivery of these proteins to the sec machinery (Perlman and Halvorson 1983). Following cleavage of the signal sequence, proteins then rely on specialised secretion machinery, comprising 12 or more proteins, and probably includes a gated channel. This pathway allows a variety of outcomes, known as terminal branches, for the secreted proteins *ie.* they may be excreted, periplasmic or envelope associated (Figure 1).

1.6.3 Type III secretion

This highly conserved export system also plays a major role in the export of virulence factors in a number of pathogens (Van Gijsegem *et al.* 1993; Hueck 1998). This secretion system is contact mediated and delivers bacterial virulence factors directly into the host cell cytoplasm which can affect host cell signalling (Kerr 1999). A large number of genes appear to be responsible for this export mechanism (approximately 20) and there is thought to be some overlap with type II export (d'Enfert 1993). This mechanism is, however, sec-independent. The prototype is the Yops (*Yersinia* outer membrane proteins) secretion pathway in *Yersinia* species (Michiels *et al.* 1990), although more recently homologues have been identifed in *B. bronchiseptica* (Yuk *et al.* 1999) and *B. pertussis* (Kerr *et al.* 1999) that are under the control of the *bvgAS*-locus and may modulate immune responses during infection by inhibiting NF-KB activation by TNFα (Yuk *et al.* 2000).

1.6.4 Type IV secretion

Recently, type IV secretion has been used to describe the autotransporter secretion system (Finlay and Falkow 1997; Henderson *et al.* 1998; Holland 1998). The prototype member is the IgA protease of *Neisseria gonorrhoeae*. Type IV secretion is

described in detail in section 1.7 Confusingly, type IV secretion has also been assigned to the totally unrelated secretion of the type utilised to mobilise DNA, or the pertussis toxin secretion pathway (section 1.8.1) (Burns 1999). This is, however, more often referred to as type V secretion.

1.6.5 Type V secretion

This mechanism involves export of an assembled protein complex through the outer membrane, in contrast to the type I-type IV secretion pathways which are presumed to transport unfolded proteins (Weiss 1994). The pertussis toxin operon of *B. pertussis* is perhaps the best described example of type V secretion and is described in more detail in section 1.8.1.

Although several mechanisms have been suggested for protein translocation across the Gram-negative cytoplasmic membrane, there is a lack of knowledge regarding the transport of outer-membrane proteins and secreted proteins through the outer membrane. There is very little evidence regarding the kinetics of the events that accompany outer membrane translocation and protein folding (unlike the inner membrane there is no ATP and no electrochemical gradient) (Nikaido and Saier 1992; Weiss 1994). Consequently, there is interest in the export, location and integration of proteins into the outer membrane. It has been suggested that targeting to the outer membrane and correct assembly may be distinct events. The precursor protein is likely to dictate such events, either at the amino or carboxy terminus or due to protein conformation.

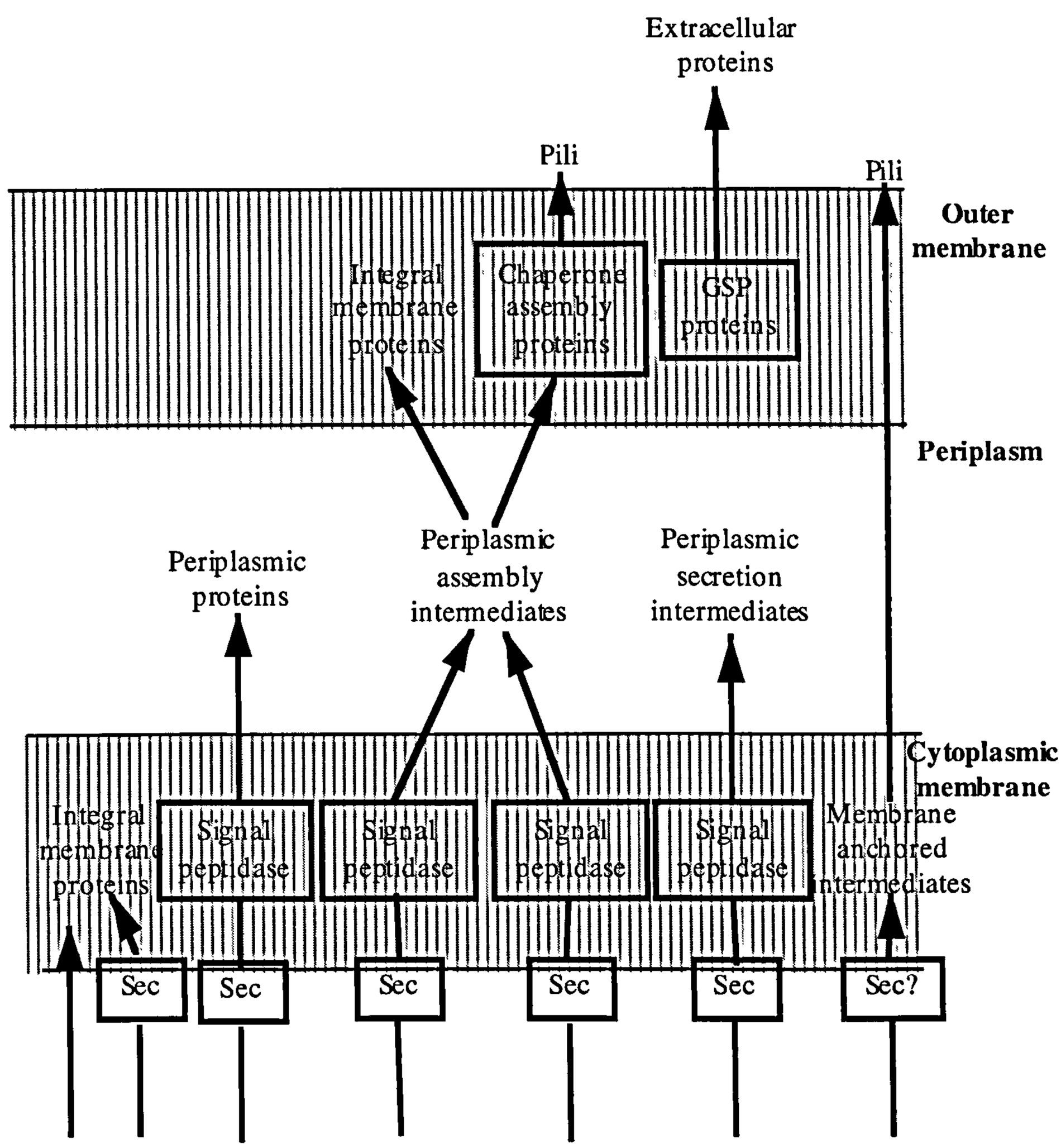


Figure 1: The main terminal branches in the general secretory pathway (GSP) of Gram-negative bacteria. Adapted from Pugsley (1993).

1.7 Autotransporter systems in Gram-negative bacteria

A newly described family of outer membrane proteins is emerging, with members present not only in *Bordetella* but also in other bacterial genera (Table 2). These autotransporters, or processed outer-membrane proteins (POMPs), have homology in their carboxy-terminus domains which are approximately 30 KDa. This conserved domain may be processed from the precursor protein and remains in the outer membrane. The amino terminal portions of these proteins have a wide variety of functions, ranging from protease activity to adhesion. The amino terminus of these proteins may be proteolytically processed and released from the cell surface, may not be processed, or may be processed yet remain cell-associated. Prior to this study, three *B. pertussis* proteins which can now be included in the autotransporter family had been described - pertactin (Prn), tracheal colonisation factor (Tcf) and *Bordetella* resistance to killing A (BrkA).

This family of processed (autotransporter) outer membrane proteins are linked by their abilitiy to traverse the outer membrane, apparently without the use of separate chaperone proteins. The putative amphipathic ß-barrel structure formed by the carboxy terminal domain is grossly consistent with the structure of bacterial porins *i.e.* an even number of antiparallel ß-sheets, with the first and last transmembrane segments in opposite directions to close the barrel (Weiss *et al.* 1991; Cowan *et al.* 1992). Models have been suggested for the structure and function of the C-terminal domains, although little has been justified experimentally. The conserved nature and abundance of these domains implies functional importance. The portion displaying the highest level of conservation is the last nine-twelve amino acids at the C-terminus (C-terminal consensus, Table 2) which consist of alternating hydrophobic amino acids and terminates in a hydrophobic, aromatic amino acid. The presence of this residue such as phenylalanine (or tryptophan) at the carboxy terminus appears to be essential for the correct assembly of these, and other, bacterial outer-membrane proteins (Stuyve *et al.* 1991). The hydrophobicity of this residue probably ensures correct orientation of the

outer-membrane protein, the reason for the aromatic character of the amino acid is less obvious (Stuyve, Moons *et al.* 1991). The region of greatest similarity between SepA, Tsh, Hsr, BrkA and Prn is at the N-terminal side of the region supposed to form ß-sheets *i.e.* immediately adjacent to the carboxy terminal domain (Benjelloun-Toulmi, Sansonetti *et al.* 1995). It is suggested that this region is essential for correct presentation of the passenger domain to the carboxy terminal pore, and in the case of IgA protease of *Neisseria gonorrhoeae*, the term linker region has been used (see Figure 2). The *N. gonorrhoeae* IgA protease is the prototype autotransporter and has extensive similarities with several recently identified proteins (Table 2) (Klauser *et al.* 1993; Benjelloun-Toulmi *et al.* 1995). Figure 2 outlines the secretion pathway proposed for the export of autotransporter proteins.

Gene	Organism	Precursor Mw (KDa)	Mature Mw (KDa)	C-terminal consensus	Major function	Ref
iga	*Neisseria gonorrhoeae*	169	106	QKSGQIKIQIRF	Protease	1
hsr	*Helicobacter mustelae*	150	NK	DYGFNIGYRYNF	Colonisation	2
ssp	*Serratia marcescens*	112	72	DNSVNAGLTWRF	Protease	3
espC	*Escherichia coli*	140	110	DNAINANFRYSF	NK	4
vacA	*Helicobacter pylori*	140	87	HFASNLGMRYSF	Cytotoxin	5
aida-1	*Escherichia coli*	132	100	NAISGALGKYSF	Adhesion	6
sepA	*Shigella flexneri*	150	110	DNAINANFRYVF	Invasion	7
icsA	*Shigella flexneri*	130	95	DTQGILGVKYTF	Intracellular spread	8
hap	*Haemophilus influenzae*	160	110	QQNVGVKLGYRW	Colonisation	9
tsh	*Escherichia coli*	148	118	DDAINANIRYSF	Haemagglutination	10
prn	*Bordetella spp.*	94	68	PWTFHAGYRYSW	Adhesion/invasion	11
brkA	*Bordetella pertussis*	103	73	PWSFHAGYRYSF	Serum resistance	12
tcf	*Bordetella pertussis*	90	68	PWTFHVGYRYAW	Colonisation	13

Table 2: Processed outer-membrane proteins. NK: Not known. References: 1 (Pohlner *et al.* 1987), 2 (O'Toole *et al.* 1994), 3 (Shimada *et al.* 1994), 4 (Stein *et al.* 1996), 5 (Cover 1996), 6 (Benz and Schmidt 1992), 7 (Benjelloun-Toulmi *et al.* 1995), 8 (Goldberg *et al.* 1993; Suzuki *et al.* 1995), 9 (St. Geme III *et al.* 1994), 10 (Provence and Curtiss III 1994), 11 (Charles *et al.* 1989), 12 (Fernandez and Weiss 1994), 13 (Finn and Stevens, 1995). Table adapted from Jose *et al* .(1995).

1.7.1 The IgA protease of *N. gonorrhoeae*

This exopolyprotein (secreted protein) exhibits much of the gross structural homology described for Prn, BrkA and Tcf (Jose *et al.* 1995). The role of the carboxy terminal domain in targeting of the protein to the outer membrane and secretion to the cell exterior has been studied (Klauser *et al.* 1992; Klauser *et al.* 1993). The term helper domain was suggested initially to describe the function of this region (Pohlner *et al.* 1987), although the term Igaß is now used (Klauser *et al.* 1992).

A three step mechanism for secretion of IgA protease has been suggested and is depicted in Figure 2. It is likely that other members of the autotransporter protein family share some of the features of this model. The first stage involves the sec-dependent targeting of the pre-pro-protein to the cytoplasmic membrane and this process is therefore dependent on a signal sequence motif at the amino terminus (Perlman and Halvorson 1983). During the second process it is assumed that the carboxy terminal domain forms a ß-barrel in the outer membrane. The tertiary structure of the protein is likely (based on computer predictions) to consist of fourteen amphipathic, antiparallel ß-sheets, forming a channel with a closed hydrophilic interior and a hydrophobic exterior (Klauser *et al.* 1992). During the final stage it is suggested that the N-terminal passenger domain is translocated to the cell exterior via the pore formed (Pohlner *et al.* 1987).

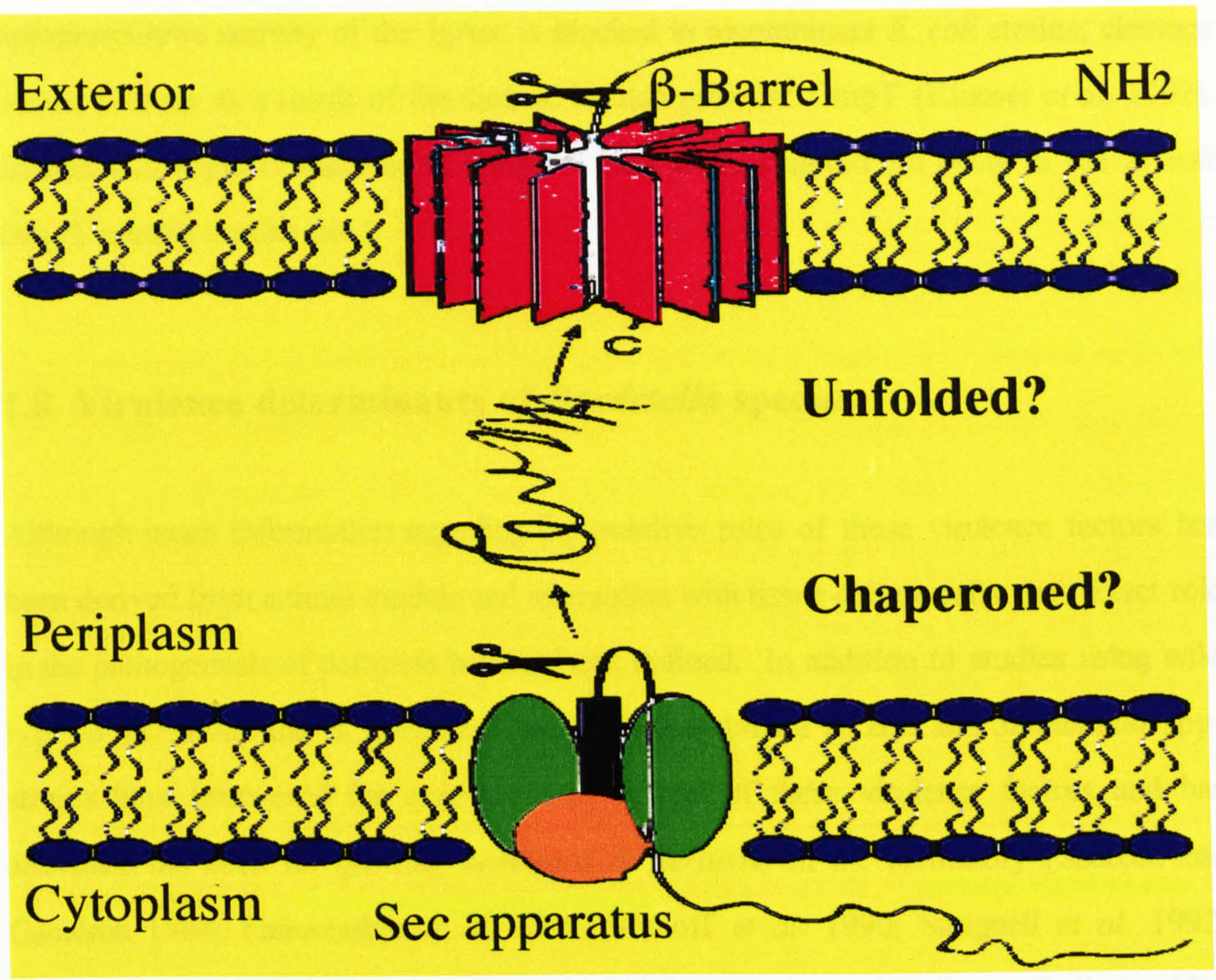


Figure 2: A model proposed for translocation of the IgA protease of *N. gonorrhoeae.* The other proteins outlined in Table 2 may also be processed in this manner.

Following translocation across the outer membrane to the cell exterior the mature protein may be released into the extracellular environment. Benz *et al.* (1992) speculated that this cleavage event (in AIDA-I) may be a essential part of the translocation across the outer membrane. Once the IgA protease reaches the exterior of the cell, autoproteolysis occurs and the mature protease is released. When the autoproteolytic activity of the IgAse is blocked in recombinant *E. coli* strains, cleavage occurs *in trans* as a result of the surface located protease OmpT (Klauser *et al.* 1992). Evidence to support this role of OmpT is that processing did not occur in the *E. coli* OmpT mutant strains used.

1.8 Virulence determinants of *Bordetella* species

Although much information regarding the putative roles of these virulence factors has been derived from animal models and interaction with tissue culture cells, their exact role in the pathogenesis of pertussis has not been defined. In addition to studies using wild type *B. pertussis* and *B. pertussis* mutants, recombinant *E. coli* and *Salmonella* spp. strains have been used for expression of several of these virulence factors and has alleviated the need for growing fastidious *B. pertussis* in the laboratory (Shareck and Cameron 1984; Fairweather *et al.* 1990; Makoff *et al.* 1990; Strugnell *et al.* 1992; Charles *et al.* 1994; Anderson *et al.* 1996). Such models have further aided characterisation of the role of individual virulence factors in *B. pertussis*. Table 1 outlines the virulence factors present in bordetellae. Not all virulence factors are present in each species. Most virulence factors have been best characterised in *B. pertussis*. A review on virulence factors has been published by Parton (1996).

1.8.1 Pertussis toxin

The most extensively studied *B. pertussis* virulence factor is pertussis toxin (PT), which has a range of biological activities, and is implicated as having a fundamental role in the pathogenesis of pertussis.

PT is composed of five subunits that are the products of five genes, which form a 105 KDa complex with ADP ribosylating activity. The A (enzymic) subunit is associated with five B subunits that are responsible for binding of the toxin to the target eukaryotic cell and insertion of the A subunit into the cytoplasm. This interaction also enables *B. pertussis* to bind to eukaryotic cells including human respiratory ciliated cells and macrophages (Sandros and Tuomanen 1993). The nature of the lectin-like interaction between the B-subunits and the eukaryotic cell mimics the relationship between the leukocyte surface and the capillary wall that occurs during early diapedesis (Sandros and Tuomanen 1993). The A subunit acts on the G-proteins of target eukaryotic cells and alters the response of these cells to external stimuli. This results in a variety of consequences including inhibition of chemotactic detection by leukocytes, lymphocytosis and increased vascular permeability (Craig *et al.* 1988). The exact role of PT in the pathogenesis of whooping cough remains obscure, although its absence in *B. parapertussis* implies that it may have only a minor role in causing the typical whooping cough symptoms (Parton 1996).

Secretion of PT is dependent on an N-terminal signal sequence on each subunit that is removed during protein maturation (Ricci *et al.* 1996). The *B. pertussis* equivalent of the sec-mediated pathway is probably responsible for transfer of these subunits to the periplasmic space (Weiss 1994). Translocation across the outer membrane is mediated by an operon of seven genes (*ptlB* to *ptlH*-pertussis toxin liberation) that have extensive homology with the outer membrane complex responsible for the transfer of DNA from the plant pathogen *Agrobacterium tumefaciens* (Weiss 1994; Ricci *et al.* 1996). In the latter system, the DNA is secreted in the form of a complex of protein that coats the

DNA. PT is also secreted as an assembled protein complex and novel machinery is needed (Weiss 1994). This export pathway has been termed type V secretion (section 1.6.5). Most proteins are presented to the translocation apparatus in the form of a long slender structure, which is secreted in an unfolded state and folds following secretion. Pertussis toxin is an exception. The export of a functional toxin complex prevents energy from being wasted on producing subunits that do not become part of a functional toxin. Secretion of protein complexes has not yet been identified in Gram-positive bacteria, probably because the outer membrane is essential in blocking the transport of unassembled subunits (Weiss 1994).

1.8.2 Adenylate cyclase toxin

Adenylate cyclase toxin (ACT) is a 200 KDa toxin which has homology to the RTX (repeats in toxin) family of proteins. This bifunctional protein is capable of haemolysis as well as adenylate cyclase activity (Wolff *et al.* 1980). Most of the toxin remains on the cell surface of the bacterium and acts as a contact toxin. Most toxins in the RTX family are pore forming, cytolytic toxins and the C-terminal portion of ACT allows the N-terminus to penetrate a wide range of eukaryotic cells and cause unregulated synthesis of cyclic AMP (Hewlett and Gordon, 1988). Immune effector cells such as monocytes, macrophages and natural killer cells are thought to be the primary targets and this toxin. The range of effects of the toxin include apoptosis and inhibition of phagocytosis, chemotaxis and microbial killing. Adenylate cyclase toxin does appear to be protective in a mouse model and may constitute a future acellular vaccine component, although fears that cross reaction to a human brain adenylate cyclase has raised doubts about its suitability (Guiso *et al.* 1989; Parton, 1998) An antigenically distinct yet functionally similar adenylate cyclase toxin is also secreted by *B. parapertussis* and *B. bronchiseptica* (Hewlett and Gordon, 1988).

1.8.3 Tracheal cytotoxin

Tracheal cytotoxin (TCT) is produced by *B. pertussis*, *B. bronchiseptica*, *B. parapertussis* and *B. avium* and belongs to the muramyl di-peptide toxin family (Goldman, 1988). TCT is an unusually small toxin (921 Da) which is a disaccharide-tetrapeptide derived from peptidoglycan which is released from the cell envelope. This toxin causes ciliostasis and cell extrusion in hamster tracheal organ cultures and inhibition of DNA synthesis in hamster trachea epithelial cultures (Goldman and Herwaldt, 1985). These effects may be due to stimulation of IL-1 production from tracheal cells which in turn stimulates nitric oxide synthesis (Heiss *et al.* 1994). TCT is also toxic for human neutrophils and may therefore contribute to persistence of the bacteria within the airways (Cundell *et al.* 1994).

1.8.4 Heat-labile toxin

Heat-labile toxin (HLT), also known as dermonecrotic toxin, is produced by *B. pertussis*, *B. bronchiseptica*, *B. parapertussis* and *B. avium* and was the first *Bordetella* toxin to be described (Iida and Okonogi, 1971; Walker and Weiss, 1994). HLT remains within the bacterial cell cytoplasm as a single polypeptide of approximately 160 KDa. Studies using *B. bronchiseptica* suggest that HLT deaminates host cell Rho which causes cytoskeletal effects (Horiguchi *et al.* 1997). *B. pertussis* mutants deficient in HLT production were unaltered in their ability to cause lethal infection in mice (Weiss and Goodwin, 1989). However, the toxin is extremely potent in a number of bioassays and exhibits potent vasoconstrictive activity on smooth muscle. Such a local inflammatory reaction on the highly vascularised tissues in the respiratory tract could account for some of the pathology of pertussis. However, the exact role of this toxin in disease has not yet been defined.

1.8.5 Endotoxin

Endotoxin, also known as lipopolysaccharide (LPS), is present in many bacterial pathogens and is a major component of the Gram-negative outer membrane. In *B. pertussis*, the O side chain is short and the LPS is referred to as lipooligosaccharide (LOS). At present there is no clear role for LOS in *B. pertussis* although it has the usual properties of endotoxins such as pyrogenicity, adjuvanticity and general toxicity. In addition, the *B. pertussis* LOS has some unusual properties in LPS resistant mice such as lymphocyte mitogenicity and polyclonal B-cell activation (Watanabe *et al.* 1990). LOS may be responsible for the mild fever in early pertussis. Much of the reactogenicity of the whole-cell vaccine preparations may be due to endotoxin.

1.8.6 Fimbriae

Fimbriae (pili; agglutinogens) are produced by *B. pertussis*, *B. bronchiseptica*, *B. parapertussis* and *B. avium*. Four fimbrial genes have been identified (*fim1,2,3* and *X*). Fim1 is present on all *B. pertussis* strains whereas Fim2 and Fim3 may be absent or expressed in various combinations on virulent cells (Livey *et al.* 1987; Robinson *et al.* 1990; Savelkoul *et al.* 1996). FimX is expressed only in *B. bronchiseptica*. Unlike the fimbriae of other bacteria the role of fimbriae in *Bordetella* infection has not been defined and, in *B. pertussis*, do not enable the bacteria to bind to human respiratory ciliated cells. A fimbrial mutant of *B. bronchiseptica* appeared to be altered in its ability to cause respiratory tract infections in rats, although adhesion to cell lines (including human laryngeal Hep-2 cells) was not affected (Mattoo *et al.* 2000). It appears that the binding of fimbriae to non-ciliated cells including monocytes may promote phagocytosis. Fimbriae on *B. pertussis* interact with very late antigen-5 (VLA-5) on monocytes which activates complement receptor 3 (CR3) on monocytes, which is a receptor for FHA (section 1.8.7). Purified fimbriae appear to confer serospecific protection against *B. pertussis* infection in mice (Willems *et al.* 1998). This is consistent with the finding

that, to some extent, serospecific protection occurs when children are immunised with whole-cell vaccine preparations (Mooi 1998).

1.8.7 Filamentous haemagglutinin

The major secreted product of *B. pertussis* is filamentous haemagglutinin (FHA), which undergoes postranslational processing of the amino and carboxy terminal domains to yield a 220 KDa secreted protein (Brown and Parker 1987; Stibitz *et al.* 1988; Domenighini *et al.* 1990; Locht *et al.* 1993; Makhov *et al.* 1994). Recently, it has been demonstrated that FHA in *B. bronchiseptica* is both necessary and sufficient for mediating adherence to a rat lung epithelial cell line, and necessary but not sufficient for tracheal colonisation in a healthy anaesthetised rat (Cotter *et al.* 1998). Cotter *et al.* (1998) suggested that the role of FHA was to overcome the clearance mechanism of the mucociliary escalator. The mature protein acts as an adhesin and forms a rigid rod based hairpin structure (Makhov *et al.* 1994). FHA contains the triplet Arg-Gly-Asp (RGD) which appears to be involved in binding of FHA to the leukocyte integrin CR3 (Relman, Wright *et al.* 1990; Sandros and Tuomanen 1993; Ishibashi *et al.* 1994). This binding may promote internalisation of *B. pertussis* into macrophages and allow persistence, as this uptake mechanism does not trigger an oxidative burst. FHA, fimbriae and PT appear to function in adhesion, as PT and fimbriae are thought to cause upregulation of the CR3 integrin on the leukocyte surface (Locht and Cabezon 1990; Sandros and Tuomanen 1993; Hazenbos *et al.* 1995).

The secondary infections often associated with pertussis may be attributable to the adhesive aspects of PT and FHA. When secreted from *B. pertussis*, other bacteria, notably *Haemophilus influenzae*, may ultilise them to adhere to human ciliated cells. This has been termed "piracy of adhesins" (Tuomanen 1988).

Carbohydrate recognition moeities have been identified within FHA that are crucial in allowing adhesion to both ciliated epithelial cells and cultured Chinese hamster ovary

cells (Domenighini *et al.* 1990). In addition, a heparin-binding domain is thought to bind glycosaminoglycans which facilitates adherence to non-ciliated cells (Sandros and Tuomanen 1993).

Targeting of the FHA precursor, FhaB, to the cytoplasmic membrane is probably performed in a sec-dependent manner and a short (22 residue) signal sequence is present (Jacob-Dubuisson *et al.* 1996). Its translocation across the outer membrane depends on an accessory protein in the outer membrane, designated FhaC (Stibitz *et al.* 1988; Domenighini, *et al.* 1990). During the maturation of FHA, the C-terminal domain of FhaB is cleaved.

The function of this cleaved carboxy terminus (150 KDa) in this instance is probably to prevent the formation of a hairpin structure within the periplasmic space, and possibly to aid presentation of the N-terminal domain to FhaC (Renauld-Mongenie *et al.* 1996). The C-terminal domain may also prevent premature interaction between the N-terminal region and FhaC during export of these proteins towards the outer membrane (Renauld-Mongenie *et al.* 1996). The C-terminal domain of FHA is probably therefore an intramolecular chaperone of FhaB (Jacob-Dubuisson *et al.* 1996). The fate of the carboxy terminal domain is unlike that of Prn, BrkA and Tcf (see below sections) and, following cleavage, is probably degraded in the periplasm. It has been predicted that the FhaB C-terminal domain anchors FHA in the cell envelope, perhaps by spanning the periplasmic space and in this way linking the inner and outer membranes (Domenighini *et al.* 1990). Proline rich repeats have been identified in the carboxy terminal domain of FhaB (Domenighini *et al.* 1990). Such a proline rich region exists in the mature Prn protein, close to the putative C-terminal processing site (Charles *et al.* 1989). These regions are thought to be characteristic of proteins that act as anchor regions and span the periplasm and therefore this may be an additional function of the FhaB C-terminus.

The proteases responsible for processing FHA are not specific to bordetellae (the same processing can occur in *E. coli*) and cleavage is likely to be the result of ubiquitous

envelope proteases (Lambert-Buisine *et al.* 198; Jacob-Dubuisson *et al.* 1996). SDS-PAGE of purified FHA results in multiple bands, all of which react with anti-FHA mAbs which suggests multiple cleavage sites (Irons *et al.* 1983; Jacob-Dubuisson *et al.* 1996). Autoproteolytic activity of FHA has not been eliminated and remains a possibility (Jacob-Dubuisson *et al.* 1996).

1.9 Autotransporter proteins of *Bordetella* species

On the basis of amino acid similarity three proteins, Prn, BrkA and Tcf have been classified as autotransporter proteins (Figure 3). The more recently described Vag 8 (section 1.13) can also be included in the autotransporter family. The receptor motifs encoded by these bordetellae adhesins are not well characterised. The tripeptide Arg-Gly-Asp (RGD) (Ruoslachti and Pierschbacher 1986; Hynes 1987) is present on FHA, Prn, Tcf and BrkA and may be important in promoting the adhesion of bordetellae to mammalian cells (Relman *et al.* 1990; Leininger *et al.* 1992). This motif is the cell attatchment site of mammalian adhesion proteins such as fibronectin, vitronectin and fibrinogen. BrkA and Tcf also contain Ser-Gly-X-Gly (SGXG) glycosaminoglycan binding motifs. Recently, the importance of proteoglycans (glycosaminoglycans attatched to a protein core), such as heparin, as co-receptors in microbial adherence and invasion have been reviewed (Rostrand and Esko 1997). It has also been suggested that binding to proteoglycans may result in host cell responses that lead to internalisation of adherent microbes (Falkow 1991).

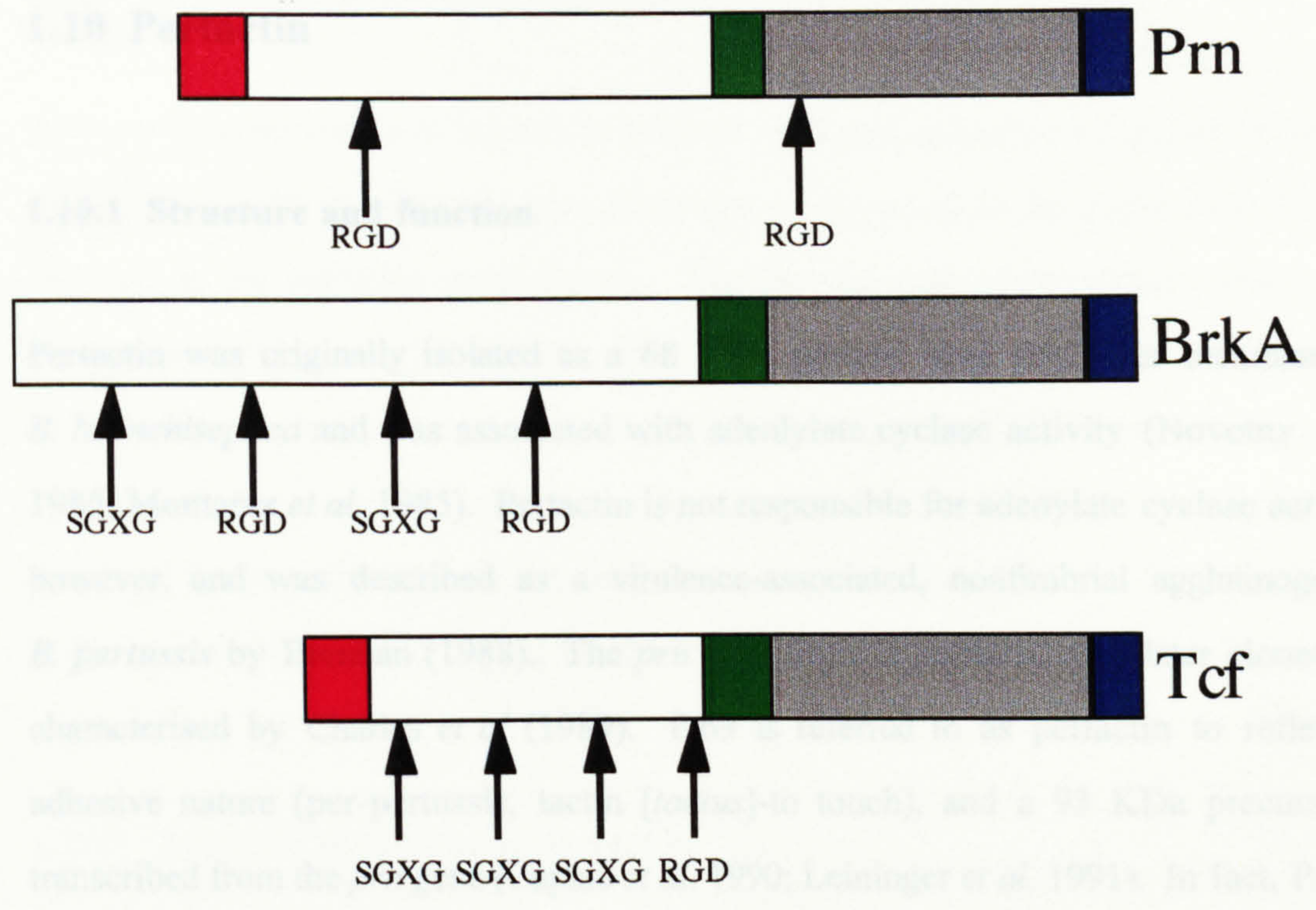
Previously, the cell envelope of virulent strains of *B. pertussis* was reported to contain at least two heat modifiable outer membrane proteins (OMPs) of approximately 30 KDa (Parton and Wardlaw 1975; Wardlaw *et al.* 1976; Ezzell *et al.* 1981; Armstrong and Parker 1986). These proteins, designated 30 KDa and 32 KDa, have since been identified as the carboxy terminal domains of BrkA and Tcf respectively (Fernandez and Weiss 1994; Finn and Stevens 1995). The sites of proteolytic processing of the *Bordetella* autotransporter proteins has been established by N-terminal sequencing

(Below) (Hamstra *et al.* 1995; Hamstra and Poolman Unpublished). It is also known that these proteins are heat modifiable and correspond to the major bands designated B and C by Parton and Wardlaw (1975) (Parton, personal communication). The processing sites of Prn, Tcf and BrkA are highlighted by // below. The main features of these proteins can be seen in Figure 3.

Prn: YAESN // ALSKRLGELRLNPDAGGAW

Tcf: QAEMN // ALSKRMGELRLTPVAGGVWGRAF

BrkA: LAESN // ALDKRLGELRLNADAGG



Signal sequence

Proteolytic processing site

Putative Beta- barrel

Outer membrane localisation signal

RGD (Arg-Gly-Asp) Integrin binding/Cell adhesion motif

SGXG Glycosaminoglycan binding motif

Figure 3: The gross structural similarities between Prn, BrkA, Tcf and Vag8.

1.10 Pertactin

1.10.1 Structure and function

Pertactin was originally isolated as a 68 KDa protein from the outer membrane of *B. bronchiseptica* and was associated with adenlylate cyclase activity (Novotny *et al.* 1985; Montaraz *et al.* 1985). Pertactin is not responsible for adenylate cyclase activity, however, and was described as a virulence-associated, nonfimbrial agglutinogen of *B. pertussis* by Brennan (1988). The *prn* gene from *B. pertussis* was later cloned and characterised by Charles *et al* (1989). P.69 is referred to as pertactin to reflect its adhesive nature (per-pertussis, tactin [*tactus*]-to touch), and a 93 KDa precursor is transcribed from the *prn* gene (Capiau *et al.* 1990; Leininger *et al.* 1991). In fact, P.69 is actually a protein of 60-61 KDa and its misleading migration on SDS-PAGE is probably a result of a high proline content (Makoff *et al.* 1990; Capiau *et al.* 1990). The crystal structure of P.69 has been determined, and the β-helical structure is the longest reported to date (Emsley *et al.* 1996).

Pertactin (P.69), in *B. pertussis*, is the representative member of the family of immunologically cross-reactive, closely-related proteins of reported molecular mass 68-70 KDa that are also present in *B. parapertussis* and *B. bronchiseptica* (Kobisch and Novotny, 1990; Charles *et al.* 1990; Li *et al.* 1991). The C-terminal domain is the most conserved region of the Prn precursor proteins. The BvgAS regulation system controls expression of the pertactin gene family. The relatively higher GC content of the P.93 gene (71.3%) could suggest that this gene evolved in a different way to other members of the *bvg*-controlled family of genes (Charles *et al.* 1994).

1.10.2 Proteolytic processing of pertactin

The *prn* gene encodes a polypeptide with an predicted molecular weight of 93478 Da (P.93). The post translational processing site is suggested to be amino acid 597-598 (lys-arg) in the pertactin proprecursor. The cleavage site of the native protein was determined by Capiau (1990) who confirmed earlier work on recombinant P.93 (Makoff *et al.* 1990). In addition there is a characteristic amino terminal signal peptide of 34 amino acids that is removed post translationally (Charles *et al.* 1989). Such a sequence is typical for proteins that are targetted through a membrane either to the periplasm or to the outside of the cell (Perlman and Halvorson 1983; Charles *et al.* 1989).

1.10.3 Adhesive properties of pertactin

Pertactin is exposed on the surface of *B. pertussis* and is a non-fimbrial antigen (Brennan *et al.* 1988; Charles *et al.* 1989). Physically longer adhesins in *B. pertussis* such as FHA or fimbriae that are capable of overcoming an electrostatic barrier probably bring bordetellae into close enough contact with the host cell to allow P.69 to act (Everest *et al.* 1996).

Adhesion of *B. pertussis* to Hep2 (human epithelium-like) cells is probably mediated, at least in part by pertactin, although such studies are complicated by the presence of other adhesins, notably FHA, that compensate in *prn*- mutant strains (Ewanowich *et al.* 1990; Leininger *et al.* 1992). However, Roberts *et al.* (1991) were able to show in a Hep-2 adhesion assay that *prn*- mutants are cabable of binding and invading cells to the same extent as wild-type *Bordetella* strains. When Chinese hamster ovary (CHO) and HeLa (Human epithelial) cell types were used for adhesion studies, P.69 mutants were seen to adhere 30-40% less and entry of *Staphylococcus aureus* into HeLa cells increased following coating with P.69 (Leininger *et al.* 1991).

The RGD integrin binding motif has been identified in mature P.69 and also in the carboxy-terminal domain (Figure 3) (Makoff *et al.* 1990). The role of the P.69 RGD is at present unclear. Althought Leininger *et al.* (1991; 1992) found that synthetic peptides containing RGD could partially inhibit bacterial adhesion to eukaryotic cells, an RGD→RGE P.69 mutant developed by Everest *et al.* (1996) did not show altered adhesive properties. Also, the 3D structure of pertactin indicates that the RGD within the N-terminus of pertactin may not be functional (Emsley *et al.* 1994; Emsley *et al* 1996). Pertactin also has two proline rich regions, $(PQP)_5$ and $(GGXXP)_5$ which are immunodominant and may be involved in fast, weak binding of P.69 to the cells of the respiratory tract (Domenighini *et al.* 1990; M. Roberts, personal communication).

It has been suggested that binding of P.69 may activate intracellular signalling pathways in eukaryotic cells and cause upregulation of the P.69 receptor. The prerequisite for metabolically-active eukaryotic cells to bind *E. coli* expressing P.69 supports this theory (Everest *et al.* 1996). Such upregulation of receptor expression is not unique to *Bordetella* species and enteropathogenic *E. coli* and *Salmonella* species are able to influence receptor expression in a similar fashion (Biska *et al.* 1993).

The C-terminal domain of pertactin, P.30, has been identified as a major component of the outer membrane of *B. pertussis* and of *E. coli* strains which have been transformed with a P.93 expression construct (Charles *et al.* 1994). There is no evidence (from cell agglutination assays and ELISA) that the C-terminal domain is surface exposed (Charles *et al.* 1994). When a truncated gene lacking the region P.30 was expressed in *B. pertussis* and *E. coli*, P.69 could not be detected at the cell surface, and appeared to accumulate in the periplasm (Charles *et al.* 1994). The role of the RGD in the P.30 domain (Figure 3) has not been evaluated.

1.10.4 Invasive properties of pertactin

There are contradictory results regarding the invasive properties of P.69 and analysis of results is often difficult as adhesion is often considered as a prerequisite for invasion. Studies using synthetic RGD-containing peptides in a competitive adhesion assay by Leininger *et al.* (1992) suggest that the RGD sequence of pertactin may play a role in entry of *B. pertussis* into HeLa cells. A *Salmonella* strain expressing P.69 showed increased invasion of CHO and Hep-2 cells , and the effect was enhanced when a rough strain was utilised (Everest *et al.* 1996). This may be due to shorter LPS in these rough strains, as seen in *Bordetella* wild types which allows better exposure of Prn. However, expression of P.69 in a non-invasive *E. coli* strain did not confer invasive properties, although adhesion was enhanced. Recent studies using human tracheal epithelial cells (HTE) suggests that pertactin inhibits invasion, perhaps by interacting with the adenylate cyclase toxin (Bassinet *et al.* 2000). The nature of this interaction has not been described. At present it appears likely that the role of pertactin in invasion is confined to adhesion.

1.10.5 Immunogenicity of pertactin

Pertactin is also immunogenic in humans as can be determined by the high level of anti-P.69 circulating and secretory antibodies in individuals that are recovering from pertussis or who have received the whole-cell vaccine (Brennan *et al.* 1988; Charles *et al.* 1989; Anwar 1991). There appears to be two repeated sequences within pertactin that are immunodominant *i.e.* are recognised by several anti-pertactin monoclonal antibodies and polyvalent antibodies (Charles *et al.* 1990). These repeat regions, $(PQP)_5$ and $(GGXXP)_5$, may serve as protective epitopes. In addition to a humoral response, several T-cell clones that recognise P.69 have been identified in an individual who suffered from pertussis during childhood (Magistris *et al.* 1986). A quarter of T-cell clones responsive to phase I *B. pertussis* recognised P.69. The possibility of using the

P.30 domain in vaccine preparations either alone or combined with other antigens has not been evaluated.

Vaccine studies in mice have uncovered the potential importance of including Prn in future vaccine formulations (Lipscombe *et al.* 1991; Roberts *et* al. 1992; Strugnell *et al.* 1992). Oral or intravenous immunnisation of mice with a live attenuated vaccine strain of *S. typhimurium* expressing P.69 induced significant protection against an aerosol *B. pertussis* challenge (Strugnell *et al.* 1992). Such vaccine strains contained the entire P.93 open reading frame and cellular responses appeared to be responsible for protection (Strugnell *et al.* 1992). Purified P.69 also conferred protection against an aerosol challenge with *B. pertussis* (Strugnell *et al.* 1992). P.69 fused to the B subunit of *E. coli* labile toxin (LTB) was immunogenic when this chimera was administered intranasally in a mouse model (Lipscombe *et al.* 1991). The lungs of mice vaccinated with this P.69-LTB fusion protein contained anti-P.69 antibodies and P.69-specific antibody-secreting cells. Mucosal immunity is considered an important aspect of protective immunity to *B. pertussis* (Lipscombe *et al.* 1991).

1.11 Tracheal colonisation factor

In 1995, a protein which contained significant homology to the pertactin carboxy-terminus was identified in *B. pertussis* but not *B. parapertussis* or *B. bronchiseptica.* (Finn and Stevens 1995). As in other autotransporters, the processing site is present and a 90 KDa form appears to be cell associated, whereas a 60 KDa form appears as a major component of culture supernatant fractions (Finn and Stevens 1995). The N-terminal domain is approximately 60 KDa and an RGD motif has been identified (Figure 3). Three SGXG glycosaminoglycan binding motifs are also present within Tcf. Therefore, Tcf may represent an adhesin which, unlike pertactin, is not anchored to the membrane and is released by proteolysis into the culture supernatant (Finn and Stevens 1995). The properties of Tcf-deficient mutants have been studied in murine models,

and have been found to show reduced ability to colonise the trachea although such strains retained the ability to survive in the lower respiratory tract (Finn and Stevens 1995). No other studies on this protein have been reported.

1.12 Bordetella resistance to killing (Brk)

The Bordetella resistance to killing locus (*brk*) encodes two separate open reading frames, *brkA* and *brkB* (Fernandez and Weiss 1994). The gene products BrkA and BrkB appear to be essential for resistance to killing by the classical (antibody-dependent) complement pathway. BrkA and BrkB also reduce the susceptibility of *B. pertussis* to antimicrobial preptides such as the defensins that are an important feature of the mucosal immune system (Fernandez and Weiss 1994; Fernandez and Weiss 1996). BrkA is similar in sequence to P.93 and, like pertactin, a role in adhesion and invasion of eukaryotic cells has been suggested (Fernandez and Weiss 1994). Unlike the Brk proteins, however, Prn appears to have no role in serum resistance of bordetellae. The BrkA protein has most homology with the Prn at its C-terminus and it contains the putative cleavage region as well as the characteristic outer membrane localisation motif. There is no amino terminal signal sequence as in the other processed outer membrane proteins and it is supposed that the protein is translocated across the cytoplasmic membrane via the BrkB gene product (Fernandez and Weiss 1994). BrkB has homology with ORFs of unknown function that have been identified in *E. coli* and *Mycobacterium leprae* (Fernandez and Weiss 1994). In addition to two RGD motifs, there are also two glycosaminoglycan (heparin) binding sites which may mimic the FHA-heparin interaction or have a role in inhibiting the polymerisation of C9. Recent evidence suggests that although some *B. bronchiseptica* strains, express BrkA, complement resistance is not mediated by this protein (Rambow *et al.* 1998).

1.13 Virulence-activated gene 8

An additional *bvg*-regulated protein may now be included in the autotransporter family on the basis of sequence similarity, virulence-activated gene 8 (Vag8) (Finn and Amsbaugh 1998). It appears that Vag8 is not processed at the outer membrane and remains as a 92 KDa form (Finn and Amsbaugh 1998). It was suggested previously that Vag8 may be important in persistence in the mouse respiratory tract as a mutant *B. pertussis* (SK8) harbouring a disrupted *vag8* gene was reduced in its ability to colonise mouse lungs (Finn *et al.* 1991). Like the other *B. pertussis* autotransporters described to date, Vag8 possesses an RGD motif. The role of this protein has not been characterised, although Vag8 mutants of *B. pertussis* were able to colonise mice as efficiently as the wild type. The authors demonstrated that this protein is the same as the 91 KDa protein decribed by Armstrong and Parker (1986) and the 92 KDa protein described by Hamstra *et al.* (1995). Hamsta *et al.* (1995) found that this protein was protective in the mouse intracerebral protection assay only when non-protective levels of pertussis toxin were present.

1.14 Protective capacity of *B. pertussis* autotransporter C-terminal domains

There is a lack of knowledge regarding the use of autotransporter C-terminal domains in vaccine preparations. Hamstra *et al.* (1995) and Hamstra and Poolman (unpublished) determined the N-terminal sequences of two proteins they were evaluating in a murine intracerebral vaccine model. It was later established that these proteins were the C-terminal portions of BrkA and Tcf. The C-terminal domains of Tcf and BrkA were purified from *B. pertussis* outer membranes and were not protective in the intracerebral mouse tests. However, the Tcf C-terminus was protective when this protein was presented as mixed protein-detergent micelles (outer-membrane complexes).

1.15 Heterologous antigen display

During the course of this study, several reviews have been published regarding heterologous antigen display (Maurer *et al.* 1997; Stahl and Uhlen 1997; Henderson *et al.* 1998). The use of heterologous antigen display using autotransporter domains has previously been highlighted using the IgA protease and AIDA-I (Klauser *et al.* 1993; Maurer 1997). Such systems have proved useful for the surface display of some proteins in *E. coli.* There is a particular interest in the construction of multivalent hybrid attenuated vaccine strains of bacteria that express protective antigens from a number of different pathogens. Additional uses for heterologous display that have been suggested include the generation of whole cell bioadsorbants for environmental purposes; as microbial biocatalysts (enzymes displayed on cell surface), or as diagnostic tools (bacteria with surface located antibody fragments) (Stahl and Uhlen 1997).

Various models have been developed which allow the export of heterologous antigens. The processing of these proteins has been investigated. Independent proteases may be responsible for the release of mature proteins in several autotransporter systems. The presence of additional proteases capable of cleaving the putative proteolysis site was confirmed using Cholera toxin B-subunit (CtxB) as a heterologous passenger protein fused to the IgA protease carboxy terminus and expressed in *E. coli* (Klauser *et al.* 1992). A feature of the autotransporter secretion system is the inability to allow passage to artificially fused domains that have stable tertiary structures, for example intramolecular disulphide bonds (Klauser *et al.* 1992). Proteins that use the autotransporter secretion system to reach the cell exterior are assumed to do so in an unfolded state and the number of cysteine residues is low (Jose *et al.* 1995). The CtxB cannot be transported using an autotransporter domain unless intramolecular disulphide bond formation can be prevented in the periplasm (Klauser *et al.* 1992). Recently, however, a single chain antibody has been fused to IgA protease beta domain and translocation of this protein, with some tertiary structure, was demonstrated (Veiga *et al.* 1999).

1.16 Aims of the project

The general aim of the investigation was to characterise members of the newly discovered family of processed outer membrane proteins of *Bordetella* species. Three virulence factors within this family had been identified *i.e.* Prn, BrkA and Tcf. During the course of this study a fourth member of this family, Vag8, was identified in *B. pertussis* (Finn and Amsbaugh 1998) but has yet to be fully characterised.

The primary aim was to identify and characterise a potentially new member of the *Bordetella* autotransporter protein family. Using primers directed towards the 3' region of the pertactin open reading frame, a similar but not identical sequence was amplified by PCR and cloned. Use of this PCR product as a probe against an existing cosmid gene library would identify the corresponding gene which could be sequenced and compared with existing sequences (Prn, BrkA,Tcf and Vag8). This C-terminus encoding region may be part of a larger precursor encoding gene, and subcloning of this gene into *E. coli* or *Salmonella* spp. for overexpression of the putative OMP may provide sufficient material for further characterisation of this protein *e.g.* as a vaccine antigen and virulence factor. Site-directed mutagenesis of the cloned OMP gene and the construction of OMP-deficient mutants would also allow an assessment of its importance in host-pathogen interactions. It would also be interesting to determine the presence of the gene in other bordetellae. Furthermore, the expression of the *bap-5* gene should be determined in these bordetellae under a variety of growth conditions.

The role of the autotransporter C-terminal fragments (ß-domains) was also to be assessed, with an emphasis on the vaccine potential of these regions in a murine model. In many studies, particularly with knockout mutants, it is possible that the functions that have been assigned to the mature protein (N-terminal product) could, in fact, be properties of the C-terminal moieties. The mechanisms of cleavage and regulation of cleavage during processing was also to be assessed along with the use of the C-terminal domains in heterologous antigen display.

2.0 Materials and Methods

2.1 General bacteriological procedures

2.1.1 Source of bacteria, storage, growth and media

All bacterial strains used in this study are given in Table 3 Unless stated, all strains were available from culture collections within the Division of Infection and Immunity, University of Glasgow. The compositions of the following media are given in Appendix I: Luria Bertani broth (LB); Luria Bertani agar (LBA); Bordet-Gengou agar (BGA); 2x Yeast Tryptone (2x YT); Casamino acids (CAA); Cyclodextrin liquid (CL); SOC and Minimal medium (MM). All media were sterilised by autoclaving at 15 p.s.i. (121°C) for 15 min exept where stated. Heat labile ingredients such as antibiotics were sterilised by filtration through a sterile 0.22 μm filter (Gelman Sciences, USA). Glassware was sterilised by heating to 160°C for 2 h. Long term storage of *E. coli* was at -70°C in LB broth supplemented with glycerol 50% (v/v). Bordetellae were resuspended in 1% CAA containing glycerol 20% (v/v) and stored at -70°C.

2.1.2 Growth of *E. coli*

E. coli strains were routinely grown overnight at 37°C on LB agar. Where necessary, 500 ml LB broth in a 2 L dimpled flask (unless otherwise stated) was inoculated from such plates and incubated overnight at 37°C with shaking at 150-200rpm. Antibiotics were added at appropriate concentrations.

2.1.3 Growth of *B. pertussis*

B. pertussis was grown on BG agar containing 15% defibrinated horse blood (E and O Laboratories, Scotland) and 10% (w/v) glycerol. Plates were placed in a humidified box and incubated for 2 -3 days at 37°C. Where necessary, a heavy inoculum was made into

500 ml of CL medium in a 2L dimpled flask which was incubated for 2-3 days at 37°C with shaking at 150-200 rpm.

2.1.4 Growth of *B. bronchiseptica*

B. bronchiseptica was grown overnight at 37°C on LB agar. Where necessary, 500 ml LB broth in a 2 L dimpled flask was inoculated from such plates and incubated overnight at 37°C with shaking at 150-200rpm.

Table 3: The bacterial strains used during this study.

Strain	Genotype	Source/Remarks
E. coli BL21(DE3)	F-*ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (λDE3)	Invitrogen; Recombinant protein expression strain
E. coli M15p(REP4)	*Nal*^{s}, *Str*^{s}, *Rif*^{s}, *Thi*$^-$, *Ara*$^+$, *Gal*$^+$, *Mtl*$^-$, *RecA*$^+$, *Uvr*$^+$, *Lon*$^+$	Qiagen; Recombinant protein expression strain
E. coli JM109	*recA*1*supE*44 *endA*1 *hsdR*17 *gyrA*96 *relA*1 *thi* Δ(*lac-pro*AB)	Promega; General transformation strain
E. coli LMG194	F-Δ*lacX*74 *galE thi rpsL* Δ*phoA* (PvuII) Δ*ara*714 *leu::Tn10*	Invitrogen. This strain is deleted for araBADC and is Str^R and Tet^R; Recombinant protein expression strain
B. pertussis Taberman	Wild-type	Isolated from a child post-mortem. Ruchill Hospital, Glasgow (1979)
B. pertussis 18-323	Wild-type	i.c. challenge strain for mouse protection tests. Manclark C. Center for Biologics Evaluation and Research, Bethedsa, MD 20892
B. pertussis BP347	Avirulent mutant. Tn insertion of Tohama. Bvg$^-$ (Vir$^-$)	Weiss, A. Molecular Genetics. University of Cincinnati
B. pertussis BBC29	Wild-type	Roberts, M. Veterinary Pathology, University of Glasgow
B. pertussis BBC30	Deletion mutant of BBC29. *Prn*$^-$	Roberts, M. Veterinary Pathology, University of Glasgow
B. pertussis 2041	Deletion mutant-brkA$^-$	Weiss, A. Molecular Genetics. University of Cincinnati
B. parapertussis	Wild-type , NCTC5952	Infection and Immunity, University of Glasgow
B. bronchiseptica	Wild-type, 214	Infection and Immunity, University of Glasgow
B. avium	Wild-type, 4480	R. Rimler, Ames, Iowa

2.2 DNA extraction

2.2.1 Genomic DNA extraction

After growth of the bacteria in appropriate broth culture, the genomic DNA was extracted according to the following method. Cells from 10 ml of stationary phase culture were harvested by centrifugation at 3,000 x *g* for 15 min in a Heraeus medifuge. The pellet was resuspended in 500 µl of 50 mM Tris, 1 mM EDTA (pH 7.5) and placed at -20°C. When frozen, 50 µl of fresh lysozyme (Sigma) solution (10 mg/ml in 0.25 M Tris, pH 8.0) was added and the solution thawed with mixing at room temperature. When just thawed the solution was placed on ice for 45 min and 100 µl of fresh step solution (Appendix I) was added. The tubes were then incubated at 50°C with occasional inversion. When the suspension cleared (after 1-5 h), 60 µl of Tris-buffered phenol (Sigma) was added and the suspension mixed for 5 min. The layers were separated by centrifugation at 19,000 x *g* in a Biofuge (Rotor HFA 14.2) for 15 min and the aqueous phase removed into a fresh microfuge tube. The phenol extraction was repeated if necessary. An ethanol precipitation was then performed using 0.1 volume of 3 M sodium acetate (pH 5.3) and 2 volumes of ethanol 100%. The DNA usually precipitated and could be spooled out using a sterile Pasteur pipette or, alternatively was centrifuged at 15,000 x *g* for 15 min in a Biofuge (Rotor HFA 14.2). The DNA was then resuspended in 100 µl of 50 mM Tris, 1 mM EDTA (pH 7.5) and left overnight at 4°C. To the solution, 20 µg of RNAse (Sigma) was added and the mixture incubated at 37°C for 2 h. The DNA was then extracted with chloroform:iso-amyl alcohol (24:1, Sigma) as for the phenol extraction. The DNA was precipitated as before and resuspended in 300 µl of sterile water.

2.2.2 Plasmid purification

Plasmid DNA was extracted using the QIAprep® Miniprep purification system (Qiagen, West Sussex, UK). Briefly, 5 ml of overnight *E. coli* was centrifuged at 10,000 x *g* for 10 min (Biofuge, Rotor: HFA 14.2). The resultant bacterial pellet was resuspended in 250 µl of Buffer P1 and lysed using 250 µl of Buffer P2 (alkali-detergent solution) for 5 min at room temperature. The macromolecules were then precipitated using 350 µl of Buffer N3 (chaotropic solution) and then centrifuged at 15,000 x g for 10 min (Biofuge, Rotor: HFA 14.2). The supernate was added to a QIAprep® column and centrifuged at 10,000 x *g* for 1 min. The flow-through was discarded and the column was then washed with 500 µl of Buffer PB (trace nuclease removal) and centrifuged at 10,000 x *g* for 1 min. The flow-through was again discarded and the final wash was performed with 750 µl of Buffer PE (containing ethanol) with a further centrifugation, as described previously. An additional centrifugation was performed following flow-through removal to ensure thorough removal of ethanol. Finally, the DNA was eluted by centrifugation after addition of 40-80 µl of sterile distilled water (preheated to 65°C).

Table 4: Plasmids used or created during this study

Plasmid Name	Comments
pGemT(30)	Amplicon MR30 AT cloned into plasmid pGEMT
pCRII(640)	*Eco*RI subclone of pGemT(30) (640 bp fragment)
pGem(PEB1)	5.2 Kb *Not*I fragment (PEB1) subcloned from cosmid 3
Cosmid 3	*B. pertussis* genomic cosmid which hybridised to P640
pCRII(640)	*Eco*RI subclone of pGemT(30) (640 bp fragment)
pCRII(PCT1)	Prn C-terminus encoding DNA (PCT1) AT cloned into pCRII
pCRII(BCT1)	BrkA C-terminus encoding DNA (BCT1) AT cloned into pCRII
pCRII(TCT1)	Tcf C-terminus encoding DNA (TCT1) AT cloned into pCRII
pCRII(Bp5CT1)	Bap-5 C-terminus encoding DNA (Bp5CT1) AT cloned into pCRII
pET11a(BCT1)	BCT1 cloned into pet11a: expression construct for BrkA C-terminus
pET33b(PCT1)	PCT1 cloned into pet33b: expression construct for Prn C-terminus
pET33b(TCT1)	TCT1 cloned into pet33b: expression construct for Tcf C-terminus
pET33b(Bp5CT1)	Bp5CT1 cloned into pet33b: expression construct for Bap-5 C-terminus
pCRII(Bp5ATG1)	Bap-5 from 1st potential ATG AT cloned into pCRII
pCRII(Bp5ATG2)	Bap-5 from 2nd potential ATG AT cloned into pCRII
pCRII(Bp5ATG3)	Bap-5 from 3rd potential ATG AT cloned into pCRII
pCRII(Bp5ATG4)	Bap-5 from 4th potential ATG AT cloned into pCRII
pQE-60(Bp5ATG1)	Bp5ATG1 from pCRII(Bp5ATG1) cloned into pQE-60: expression construct of Bap-5 (from 1st potential ATG)
pQE-60(Bp5ATG2)	Bp5ATG2 from pCRII(Bp5ATG2) cloned into pQE-60: expression construct of Bap-5 (from 2nd potential ATG)
pQE-60(Bp5ATG3)	Bp5ATG3 from pCRII(Bp5ATG3) cloned into pQE-60: expression construct of Bap-5 (from 3rd potential ATG)
pQE-60(Bp5ATG4)	Bp5ATG4 from pCRII(Bp5ATG4) cloned into pQE-60: expression construct of Bap-5 (from 4th potential ATG)

Plasmid Name	Comments
pCRII(Kana)	Kanamycin resistance cassette amplified from pUC4K AT cloned into pCRII
pGemT(Kana)	Kanamycin resistance cassette amplified from pUC4K AT cloned into pGemT
pCR-Script(Bap-5K^r)	*bap-5* disrupted with Kanamycin
pCRII(PCTlink)	Prn C-terminus and linker-encoding DNA (PCTlink) AT cloned into pCRII
pBAD/gIIIA(PCTlink)	PCTlink from pCRII(PCTlink) cloned into pBAD/gIIIA: expression construct of Prn C-terminus and linker fused to a signal sequence
pCRII(NTS)	Bap-5 specific portion (NTS) AT cloned into pCRII
pQE-60(NTS)	Bap-5 specific portion (NTS) cloned into pQE-60: expression construct of NTS

2.3 Agarose gel electrophoresis

2.3.1 Sample preparation

The sample DNA (5-30 µl) was mixed with 6x DNA loading buffer (Appendix I) in a vol:vol ratio of 5:1 prior to loading into the wells. Molecular weight markers (1 Kb ladder, Gibco BRL, UK) were utilised according to manufacturer's instructions.

2.3.2 Gel preparation

Agarose (type II-A medium EEO, Sigma) was suspended in 0.5x Tris-borate-EDTA (TBE) buffer (Appendix I) at a suitable concentration and heated until the agarose was completely dissolved. The solution was allowed to cool (hand-hot) and ethidium bromide (Bio-Rad, UK) was added to a final concentration of 0.5 µg/ml. A gel tray was prepared by taping the edges with adhesive tape and the gel cast to the desired thickness. Upon setting, the gel was immersed in 0.5x TBE containing ethidium

bromide (0.5 µg/ml) in a horizontal submarine electrophoresis tank (E-C Apparatus Corporation, USA).

2.3.3 Electrophoresis

A powerpack (model SL3655, Scotlab, UK) was used to provide a constant voltage corresponding to 1-5 volts/cm. Electrophoresis was carried out until the marker dye migrated an appropriate distance.

2.3.4 Visualisation of DNA

A UV transilluminator (model TM-40, UVP Inc., California, USA) was used to visualise the ethidium bromide-stained DNA. Images were stored electronically as appropriate using the Ultra Violet Products Gel Documentation System-Image Store 5000, version 7.2 (Ultra Violet Products, Cambridge, UK). Images were printed using a video graphic printer (model UP-860, Sony). Where appropriate, photographs were captured with Fotolook (version 2.07.2 Agfa, UK) using a flatbed scanner (model Studioscan II*si*, Agfa). Electronic images were edited using Adobe Photoshop 3.0 and images labelled with Microsoft Powerpoint 4.0

2.3.5 Gel extraction procedure

DNA was purified from agarose gels using the Qiaex®II purification kit (Qiagen) according to manufacturer's instruction. The band of interest was excised from the gel and solubilised in Buffer QX1 (usually 3 volumes of buffer to one volume of gel). Qiaex resin (10-15 µl) was introduced and the mixture incubated at 50°C for 10 min. Following centrifugation for 1 min at 10,000 x *g* the supernate was removed. The pellet was then washed once with 500 µl of Buffer QX1 and twice with 500 µl of Buffer PE,

with centrifugation as before. The resin/DNA pellet was air-dried, 5-25 µl of sterile distilled water was added and, after incubation for 5 min at 50°C, the resin was removed by centrifugation as before and the eluted DNA was retained.

2.4 Estimation of DNA concentration

The concentration of DNA was estimated by measuring the absorbance at 260nm in a 1cm quartz cuvette and by applying the equation:

Concentration of DNA (ng/µl) = A_{260nm} x 50 x dilution factor

Alternatively, band intensity of a sample of DNA was compared with those of molecular weight standards of known concentration after agarose gel electrophoresis.

2.5 Concentration of DNA

To precipitate and concentrate the DNA, 3 volumes of ethanol 95% (v/v) and 0.1 volume of 3 M sodium acetate (pH 5.3) was added to the sample and the resultant mix stored at either -20°C overnight or at -80°C for 30 min depending on the time available. The sample was then centrifuged at 19,000 x *g* (Biofuge, Rotor: HFA 14.2) for 15 min and the supernate discarded. The DNA pellet was then washed with cold ethanol 70% (v/v) and dried at room temperature for 10-15 min. The DNA was resuspended in an appropriate volume of sterile distilled water and if necessary heated to 65°C for 30 min to aid dissolving.

2.6 Restriction enzyme reactions

All reactions were performed according to manufacturers' instructions (Gibco BRL, New England Biolabs or Promega) with the exception of Promega universal reaction buffer occasionally replacing the suggested buffers for simultaneous digestions with

more than one restriction enzyme. Total reaction volumes of 30 µl and 50 µl were commonly used. Where appropriate, the enzymes were heat inactivated (65°C for 15 min) and the DNA was purified using the Qiaex®II kit, often in conjunction with agarose gel electrophoresis if isolation of a single restriction product was desirable.

2.7 Polymerase chain reaction

Optimisation of the PCR conditions, where necessary, was performed according to the orthogonal array method described by Cobb *et al.* (1994). The annealing and elongation thermal parameters were then varied to obtain optimal conditions. A Hybaid thermal cycler (model: Touchdown, Hybaid Ltd, Middlesex, UK) was used for all reactions. The products of the PCR were stored at -20°C or used immediately.

2.7.1 Primers

During primer design, care was taken to avoid potential internal secondary structure and where possible a GC clamp was engineered at the 3' end of the primer. Additionally, for PCR, primers were checked so as to avoid overlap and possible dimerisation, potential for secondary structure and for compatible Tm values. Tm values were calculated according to the equation:

$$Tm(°C) = 4(G + C) + 2(A + T) - 5°C$$

Primers (50 nmol, desalted and deprotected) were obtained from Gibco BRL (Paisley, UK) and resuspended in sterile distilled water to give final concentrations of 50 pmol/µl or 0.5 pmol/µl for PCR and automated sequencing respectively. Table 5 shows the primers used in this study.

Table 5: Primers used during this study

Name	Sequence	Function
PEB1FOR1	GTT TTC CCA GTC ACG AC	Forward sequencing primer (figure 7)
PEB1FOR2	CTA TCT GGG CGA AGT CGA GG	Forward sequencing primer (figure 7)
PEB1FOR3	GCA GGT ACA TCA GCG ACA GG	Forward sequencing primer (figure 7)
PEB1FOR4	TGC CAT AGG AGC GCA GGC	Forward sequencing primer (figure 7)
PEB1FOR5	GAC TCG GTC ATS AAC GCG CTG G	Forward sequencing primer (figure 7)
PEB1FOR6	GAG CTC GCA TCG TAC GTA CCG	Forward sequencing primer (figure 7)
PEB1FOR7	CTG TAC CTG GA A AGC CTG GC	Forward sequencing primer (figure 7)
PEB1FOR8	GAT TGA TCG AAG CCA GTC TGT AGC	Forward sequencing primer (figure 7)
PEB1FOR9	CGGTA TCGGGATATATGG	Forward sequencing primer (figure 7)
PEB1FOR10	AGC ACG ATA GAT GCA TAC GGC	Forward sequencing primer (figure 7)
PEB1FOR11	GCG CAGGTCGATA TGAAAGG	Forward sequencing primer (figure 7)
PEB1FOR12	GTT GTT GCC CAT GAC GCT GC	Forward sequencing primer (figure 7)
PEB1FOR13	CAG CAC AAG GTG CTG GTG C	Forward sequencing primer (figure 7)
PEB1FOR14	CAA GGT GCT GGT GCG AGG AGC	Forward sequencing primer (figure 7)
PEB1FOR16	AAC GGA CCA GCC AGC TTT CG	Forward sequencing primer (figure 7)
PEB1FOR17	TTT CGA TTC GAC CCT GCG	Forward sequencing primer (figure 7)
PEB1REV1	CAG CTC AGG GTG GCA TAG G	Reverse sequencing primer (figure 7)
PEB1REV2	CGT GAA GTC GTT CTC GAA GC	Reverse sequencing primer (figure 7)
PEB1REV3	CCT TGC TGT C CT GCG ATG GC	Reverse sequencing primer (figure 7)
PEB1REV4	CTT GTT GTC GAG CTG CTG C	Reverse sequencing primer (figure 7)
PEB1REV5	TTG TGC TGM CCG MTG GCC TCG	Reverse sequencing primer (figure 7)
PEB1REV6	CCA TGA CCC GCG CAT TCG CC	Reverse sequencing primer (figure 7)
PEB1REV8	GCA ATG GCG GCG ACC TCT CC	Reverse sequencing primer (figure 7)
PEB1REV9	CAG CGA CGT TAT TTG CAC CC	Reverse sequencing primer (figure 7)
PEB1REV11	CTC GAT ATT CAG CGT ACC G	Reverse sequencing primer (figure 7)
PEB1REV12	TGT AAA TGT CCG CAT CCG ACA GAG C	Reverse sequencing primer (figure 7)
PEB1REV13	TGT CAT TCA TAT TCA AAT CGG	Reverse sequencing primer (figure 7)
GSP1	CGA CCA TGC CAT CGC AGG ACA GCA AG	Gene-specific primer for GenomeWalker™ primary PCR (section 3.3.1.6.3)
GSP2	GCG CAA GGC GCA GCT TCA TCG ATG AC	Gene-specific primer for GenomeWalker™ secondary PCR (section 3.3.1.6.3)
AP1	GTA ATA CGA CTC ACT ATA GGG C	Adaptor-specific primer for GenomeWalker™ primary PCR (section 3.3.1.6.3)
AP2	ACT ATA GGG CAC GCG TGC T	Adaptor-specific primer for GenomeWalker™ secondary PCR (section 3.3.1.6.3)

Name	Sequence	Function
KANANCOI	CCA TGG CCG TCG ACC TGC AGG	Primer used to amplify the kanamycin cassette from pUC4K (anneals to both sides of the cassette)(section 3.4)
BRKAC-TERM5'	TTT CCA GCA TCC ATA TGG CCG AGT CC	Expression primer BrkA C-term 5' (section 3.2)
BRKA 3'	TAT GGA TCC TTG GAG CTC GCT CAG	Expression primer BrkA C-term 3' (section 3.2)
PRNC-TERM5'	ACG CCG AAC ATA TGG CGT TGT CCA AGC	Expression primer Prn C-term 5' (section 3.2)
PRN3'	ATT GGA TCC TTT ACC AGC TGT ACC GG	Expression primer Prn C-term 3' (section 3.2)
TCF5'C-TERM5'	CCA TGG CG TTG AGC AAG CGC ATG GGC	Expression primer Tcf C-term 5' (section 3.2)
TCF3'	ATT GGA TCC CTA CCA GGC GTA GCG	Expression primer Tcf C-term 3' (section 3.2)
Bap-5CTERM5'	GAA CAT ATG GCG CTC TCC AAG CG	Expression primer Bap-5 C-term 5' (section 3.2)
Bap-5CTERM3'	GGA TCC CTA CCA GGT GTA GCG ATA GCC	Expression primer Bap-5 C-term 3' (section 3.2)
NTM1EXP5'	CCG TCA TGA GCG ACA CCT GC	5' Expression primer for Bap-5 expression from 1st potential methionine start codon (section 3.2)
NTM2EXP5'	GTC ATG AGT GCA AAT AAC GTC G	5' Expression primer for Bap-5 expression from 2nd potential methionine start codon (section 3.2)
NTM3EXP5'	GTC ATG AAC GCC ACG TTC GG	5' Expression primer for Bap-5 expression from 3rd potential methionine start codon (section 3.2)
NTM4EXP5'	GTC ATG AAA GGC GGG CGC ATT CTG GC	5' Expression primer for Bap-5 expression from 4th potential methionine start codon (section 3.2)
NTMEXP3'	AGA TCT CTA CCA GGT GTA GCG ATA GCC	3' Expression primer for all Bap-5 expression constructs (section 3.2)
NTSFOR	CCA TGG GTG CAA ATA ACG TCG CTG TGG	5' primer for expression of Bap-t specific portion, NTS (section 3.2)
NTSREV	AGA TCT CAA GGC GTA GTC GTA GTA ATA CTC	3' primer for expression of Bap-t specific portion, NTS (section 3.2)
BADPCT5'	GAG TCT GCC GCC AAC GGC AAT GGG	5' primer for amplification of the Prn C-terminus and linker region for expression (section 3.3)
BADPCT3'	GAA TTC TTA CCA GCT GTA CCG GTA GCC	3' primer for amplification of the Prn C-terminus and linker region for expression (section 3.3)
BAP5MUT5'	GCA GAG ATC TTG ATG GCA CCT CG	5' primer for amplification of Bap-5 during mutant construction (section 3.4)
BAP5MUT3'	GTA GCG AGA TCT CAG GTG GAA CGT CC	3' primer for amplification of Bap-5 during mutant construction (section 3.4)
M13FOR	GTT TTC CCA GTC AGT CAC GAC	Standard sequencing primer
M13REV	CAG GAA ACA GCT ATG AC	Standard sequencing primer

2.7.2 HotStarTaq^{TM} method

Hot-start PCR was performed according to the HotStarTaq^{TM} PCR kit (Qiagen) according to manufacturer's instruction. The following mastermix was prepared in a thin walled 0.5 ml tube immediately before use (per reaction):

10X PCR Buffer (containing 15 mM $MgCl_2$)	10 µl
dNTP mix (10 mM of each)	2 µl
5' primer (50 pmol/µl)	1 µl
3' primer (50 pmol/µl)	1 µl
HotstarTaq DNA Polymerase	0.5 µl
Template DNA	≤1 µg
Distilled water	Variable
Total volume	**100 µl**

In addition to the above components, Q-Solution was often added. Q-solution changes the melting behaviour of the DNA and is particularly useful when amplifying GC rich templates. Q-solution was supplied at 5x concentration and the volume of distilled water adjusted accordingly.

The following thermocycling parameters were used:

Initial activation step:	15 min	95°C
30 Cycles of		
Denaturation:	45 s	95°C
Annealing:	45 s	50-68°C
Approximately 5°C below T_m of primers		
Extension:	1 min	72°C

For PCR products longer than 1 Kb the extension time was increased by 1 min per Kb DNA

Final extension step of 72°C for 15 min.

2.7.3 Expand™ high fidelity PCR system

Two mastermixes were prepared according to the manufacturer's instruction (Boehringer Mannheim, Germany). A total volume of 100 µl was routinely used regardless of the predicted amplimer size.

Mastermix 1:

dNTP mix (10mM)	4 µl
Upstream primer (50pmol/µl)	1 µl
Downstream primer (50pmol/µl)	1 µl
Template DNA	variable
dH_2O	variable
Total volume	**50 µl**

Mastermix 2:

10X Expand™ High Fidelity buffer (containing 15mM $MgCl_2$)	10 µl
Expand™ High Fidelity PCR System enzyme mix	0.75 µl (2.6U)
dH_2O	39.25
Total volume	**50 µl**

The above mastermixes were combined in a 0.5 ml thin walled tube and thermal cycling was performed as described by the manufacturer except that targets larger than 1 Kb were amplified using an extension temperature of 68°C and no initial activation step was performed.

2.8 Universal GenomeWalker™ procedure

Libraries were constructed according to the Universal GenomeWalker™ kit user manual (Clontech, USA), except that digested genomic DNA was purified according to the Qiaex®II method in place of the suggested phenol-chloroform extraction.

2.8.1 Preparation of genomic DNA

Chromosomal DNA was extracted from *B. pertussis* (strain Taberman) according to section 2.2.1 and five restriction digests (with *Dra*I, *Eco*RV, *Pvu*II, *Sca*I and *Stu*I) were set up seperately using the following protocol:

Genomic DNA (0.1 µg/µl)	25 µl
Restriction enzyme (10 U/µl)	8 µl
Restriction enzyme buffer	10 µl
Deionised distilled H_2O	57 µl

The above were mixed gently and incubated at 37°C for 2 h. The reactions were vortexed at slow speed for 5-10 s and returned to 37°C for 16-18 h. Agarose gel electrophoresis was performed to examine 5 µl of each reaction for complete digestion. The remaining DNA then was purified using the Qiaex®II method (section) and resuspended in 20 µl of TE.

2.8.2 Ligation of genomic DNA to GenomeWalker adaptors

From each tube of purified, digested DNA 4 µl was added to the following reaction:

GenomeWalker Adaptor (25 µM)	1.9 µl
5x ligation buffer	1.6 µl
T4 DNA ligase (1 U/µl)	0.5 µl

The above reaction was incubated at 16°C overnight. The reactions were stopped by incubation at 70°C for 5 min with 72 µl of TE and the tubes vortexed at slow speed for 15 s.

2.8.3 PCR-based DNA walking in GenomeWalker libraries

2.8.3.1 Primary PCR

A nested PCR was performed in a 0.5 ml thin-walled tube using the two gene specific primers, GSP1 and GSP2 (table 5 and figure 8) and two adaptor specific primers, AP1 and AP2 (table 5). A primary PCR mastermix was prepared (per reaction):

	µl
Deionised water	37.8
10X *Tth* PCR reaction buffer	5
dNTP (10 mM each)	1
Mg(acetate) (25 mM)	2.2
AP1 (10 µM)	1
GSP1 (10 µM)	1
Advantage *Tth* Polymerase Mix (50X)	1

To each reaction 1 µl of each GenomeWalker DNA library was added and the following thermocycling parameters used:

STAGE 1	**7 cycles:**		
		94°C	25 s
		72°C	3 min
STAGE 2	**32 cycles:**		
		94°C	25 s
		67°C	3 min
STAGE 3	**1 cycle:**		
		67°C	7 min

2.8.3.2 Secondary PCR

The secondary PCR was performed under the same conditions as the primary PCR with the following exceptions: The template for amplification was 1 µl of the primary PCR reaction mixture. Primers AP2 and GSP2 were used (in the place of AP1 and GSP1). Stage 2 was replaced with 20 cycles of 94°C for 25 s and 67°C for 3 min.

From each reaction, 8 µl was analysed by agarose gel electrophoresis.

2.9 Cloning

2.9.1 PCR product cloning

2.9.1.1 Cloning into pCR®2.1-TOPO

PCR products were routinely cloned using the pCR®2.1-TOPO vector (also referred to as pCRII in future chapters) included in the TOPO TA cloning® kit (Invitrogen, The Netherlands). Amplicons of interest were excised following agarose gel analysis and DNA extracted using the Qiaex®II kit. The DNA was resuspended to a concentration of approximately 50 ng/µl. Ligation of the PCR product into pCR®2.1-TOPO was performed according to the manufacturer's instructions. Briefly, 4 µl of cleaned PCR product was mixed gently with 1 µl of TOPO mix (containing topoisomerase, prepared pCR®2.1-TOPO and ligation buffer). After incubation for 5 min at room temperature, the mixture was put onto ice. Immediately, one vial of TOP10F' One Shot™ competent cells were thawed gently on ice and 2 µl of mercaptoethanol (0.5 M) was added. To the cells, 2 µl of the ligation mix was added with gentle mixing. After incubation on ice for 30 min, the cells were then heat-shocked at 42°C for 30 s and returned onto ice for 2 min. SOC medium (250 µl; Appendix I) was added and the cells shaken at 37°C for 2 h. Recombinants were selected according to section 2.10.3.

2.9.1.2 Cloning into pCR-SCRIPT™

Cloning was performed using the pCR-SCRIPT™ cloning kit (Stratagene, Cambridge, UK) according to the manufacturer's instructions. This procedure involves blunt cloning of the 'polished' PCR product into a predigested *Srf*I vector site. The ligation occurs in the presence of *Srf*I which reduces the frequency of vector self-self ligations.

PCR products were electrophoresed in agarose gel, purified by gel extraction (Qiaex®II) and then 'polished' to generate blunt ends. The following were assembled, in order, in a 0.5 ml tube and incubated at 72°C for 30 min:

	µl
PCR product	10
10 mM dNTP mix (2.5 mM each)	1
10X polishing buffer	1.3
Pfu DNA polymerase (0.5 U)	1

The 'polished' PCR product was then added to the ligation reaction containing the following:

	µl
pCR-Script™ Amp SK(+) cloning vector (10 ng/µl)	1
PCR-Script 10X reaction buffer	1
ATP(10 mM)	0.5
Polished PCR product	2-4
*Srf*I (5 U/µl)	1
T4 DNA ligase (4 U/µl)	1

The ligation mix was mixed gently and incubated at room temperature for 1 h, then incubated at 65°C for 10 min, transformed into *E. coli* and a blue-white screen was performed (section 2.10.3).

2.9.2 Standard cloning protocol

2.9.2.1 DNA preparation

Vector and insert DNA were subjected to restriction endonuclease digestion to obtain complimentary cohesive overhangs and the resultant fragments were gel purified. DNA concentrations were estimated according to section 2.4 and ratios determined as follows:

$$\frac{\text{100 ng vector X ?Kb Insert}}{\text{?Kb vector}} = \text{? ng insert DNA to every 100 ng vector for a 1:1 ratio}$$

Insert:Vector ratios of 1:1 and 3:1 were commonly used.

2.9.2.2 Ligation strategies

Ligation reactions were performed in a total volume of 10 µl using 3 units of T4 DNA ligase (Promega, USA) and Promega ligase buffer. After incubation at 16°C for 18 h the reactions were transferred to 4°C for 2 h. The ligase was inactivated by heating to 70°C for 10 min and the products stored at -20°C until use.

2.9.2.3 Ligation Express™ Kit

The ligation was performed according to the manufacturer's instructions (Clontech). Compatible vector and insert termini were prepared and ratios calculated as described previously (section 2.9.2.1). Approximately 300 ng of total DNA was used per reaction (maximum volume of DNA solution was 7 µl). To the DNA 10 µl of plasmid ligation buffer, 1.2 µl of ATP (10 mM) and 1 µl of T4 DNA ligase (100 U) were added. The total volume was 13.2-20 µl. After thorough mixing the reaction was incubated for 30

min at 16°C. To this, 1 µl of glycogen (20 µg), 85 µl of distilled water and 250 µl of ethanol 100% were added. The DNA was precipitated using a dry ice/ethanol bath for 30-60 min and the DNA recovered by centrifuging at 19,000 x *g* in a Biofuge (Rotor HFA 14.2). After careful removal of the supernate the pellet was briefly air-dried and a wash in ethanol 70% (v/v) was performed to remove salts. The DNA was resuspended in 5 µl of distilled water and stored at -20°C.

2.10 Standard transformation

2.10.1 Preparation of electroporation-competent cells

An overnight culture of *E. coli* was diluted 1 in 100 in 500 ml of LB in a 2-L dimpled flask. The flask was shaken vigorously at 37°C until an ODλ600nm of 0.6-1.0 was obtained. The flask was then chilled on ice and cells harvested at 5,000 x *g* for 15 min in a Sorvall superspeed (rotor GS-3) . The resultant bacterial pellet was resuspended in 500 ml of cold sterile distilled water and centrifuged as described previously. The cell pellet was then resuspended in 250 ml of cold sterile water, centrifuged as before and resuspended in 10 ml of cold sterile glycerol 10% (v/v). Following a final centrifugation the cells were resuspended in 1.5 ml of cold sterile glycerol 10% (v/v) and 100 µl aliquots were snap frozen in liquid nitrogen. Cells remained cold (on ice) at all times during the procedure and were stored at -70°C.

2.10.2 Electroporation procedure

Prior to electroporation, a 0.2 cm electroporation cuvette (Flowgen, UK)) and the safety chamber were chilled at -20°C. To the cold cuvette, 1-5 µl of DNA (ligation mix or plasmid preparation) and 40 µl of competent cells were added. The mixture was shaken to the bottom of the cuvette and a Bio-Rad Gene Pulser (model 1652078, Bio-Rad

laboratories, UK) set at 2.5 KV, 25 μFD was connected to a Bio-Rad pulse controller (model 1652098) set to 200 Ω. The cells were pulsed once to give a time constant of 4-5 msec. Immediately following electroporation, 1 ml of prewarmed SOC medium was added and the cells incubated at 37°C without shaking for 1 h.

2.10.3 Blue-white screening of recombinants

Following incubation of the transformed cells in SOC, 80 μl was spread onto pre-warmed LB agar plates containing appropriate antibiotics, 2 μg/ml of isopropyl β-D-thiogalactopyranoside (IPTG) (Sigma) and 40 μg/ml of 5-bromo-4 chloro-3-indolyl-β-D-galactodisase (X-GAL) (Sigma). To allow recovery of recombinants from low efficiency transformations, the remaining SOC was centrifuged and the cell pellet was resuspended into 150 μl of SOC and spread onto another plate. The plates were then incubated at 37°C overnight. Resultant white colonies (potentially containing insert), due to disruption of vector encoded β-galactosidase, were cultured and plasmid DNA prepared according to section 2.2.2. Plasmid DNA was analysed by endonuclease digestion (section 2.6) for the presence of an insert.

2.11 Southern blotting

The composition of all solutions is described in Appendix I. Unless otherwise stated all reagents were supplied by Boehringer Mannheim. The samples to be analysed were electrophoresed on an ethidium bromide-stained agarose gel. The gel was then floated in denaturing solution for 45 min, followed by soaking for 30 min in neutralising solution. The gel was then blotted overnight by capillary action onto nylon membrane (Hybond N+, Amersham Pharmacia Biotech, UK). After blotting overnight, the nylon membrane was removed and the DNA was crosslinked to the membrane using a UV crosslinker (Spectrolinker XL-1000, Spectroline, New York, USA). The membrane was then

prehybridised for a minimum of 4 h in a rolling hybridisation oven (model HB-1D, Techne, Cambridge, UK) with at least 20 ml of prehydridising solution. The temperature of the oven was set within the range 50°C and 68°C depending on the desired specificity. This temperature was maintained for the subsequent hybridisation and washing steps (see below).

2.11.1 Preparation of digoxigenin-labelled probes

Template DNA was excised from an agarose gel with a sterile scalpel. Extraction was done by the Qiaex®II method and the concentration of DNA estimated as described in section 2.4. The random priming method was utilised as described in The DIG System User's Guide for Filter Hybridisation (Boehringer Mannheim). Briefly, approximately 1 µg of template DNA was diluted in distilled water to a total volume of 15 µl. The DNA was denatured in a boiling water bath for 10 min followed by immediate chilling on ice. Hexanucleotide mixture (10x) and dNTP labelling mixture (10x), 2 µl of each, were added to the tube on ice. To the reaction, 1 µl of labelling-grade Klenow DNA polymerase was added to a final concentration of 100 U/ml. After brief centrifugation, the reaction was incubated overnight at 37°C, then terminated by adding 2 µl of EDTA (200 mM, pH 8). The contents of the labelling tube were then added to 8 ml of prehybridisation solution and stored at -20°C. Denaturation of the probe was performed for 10 min in a boiling water bath followed by immediate chilling on ice.

2.11.2 Probe hybridisation

The prehybridising solution was discarded and the denatured probe immediately added to the roller tube. Hybridisation was performed for a minimum of 8 h. The membrane was washed at the hybridisation temperature for 2 x 15 min in 50 ml of washing buffer 1 followed by 2 x 15 min washes in 50 ml washing buffer 2.

2.11.3 Chemiluminescence detection

Detection was performed according to the protocol described in The DIG system User's Guide for Filter Hybridisation (Boehringer Mannheim). The washed membrane was equilibrated for 1-5 min in wash buffer. Antibody blocking solution was then incubated with the membrane for 30 min. The antibody (anti-DIG conjugated to alkaline phosphatase) was then diluted to 1 in 10,000 in the antibody blocking buffer and the membrane was then incubated for 30 min in at least 20 ml of antibody solution. The membrane was then washed for 2 x 15 min in washing buffer and then equilibrated in detection buffer for 5 min. The alkaline phosphatase substrate used (CSPD®) was diluted to 1 in 100 min detection buffer and this was dripped onto the membrane prior to sealing within a plastic sheet, for 5 min. The excess substrate was removed using a damp paper towel on the top plastic surface. The plastic was heat sealed and then incubated at 37°C for 15-45 min. The membrane was then exposed to photosensitive film (Kodak, UK) for up to 24 h, although 1-4 h was usually sufficient. After exposure the autoradiograph was immersed in developing solution (LX24, Kodak, UK) with agitation for 3 min. The film was then rinsed in water and fixed, with agitation, in Unifix (Kodak, UK) for 15 min.

2.12 Automated DNA sequencing

2.12.1 Preparation of sequencing reactions

Plasmid and cosmid DNA were purified using the QIAprep® Miniprep purification system (section 2.2.2) and reaction mixtures were assembled in thin-walled 0.5 ml tubes.

Terminator Ready Reaction Mix (Perkin-Elmer, UK):	8 µl
Double Stranded DNA:	250-500 ng
Primer	3.2 pmol
Distilled water	variable
Final Reaction Volume	**20 µl**

Where necessary, 5% DMSO was included for GC-rich templates.

2.12.2 Cycle sequencing

The above reaction mixture was placed into a thermal cycler and the following parameters used:

30 cycles of

96°C	30 s
60°C	4 min

2.12.3 Purification of extension products

Fluorescently-labelled extension products were recovered by adding 50 µl of ethanol 95% (v/v) and 2 µl of 2M sodium acetate (pH 4.6). The tubes were mixed by hand and incubated on ice for 15 min. The tubes were then centrifuged at 19,000 x *g* for 15 min using a Biofuge (Rotor HFA 14.2). The ethanol was aspirated and the pellet washed with 250 µl of ethanol 70% (v/v) without centrfugation. The pellet was then dried at room temperature for 15 min and stored at -20°C.

2.12.4 Running the sequencing gel

An Applied Biosystems 377 or 373 stretch automated sequencing machine (Perkin Elmer, UK) was used for all sequencing reactions. The sequencing service at the Molecular Biology Support Unit (MBSU) at the University of Glasgow performed the gel run and results were posted electronically. Briefly, the dried pellet was resuspended by vortexing in 4 µl of loading solution (Perkin-Elmer) containing deionised formamide and 50 mM EDTA in the ratio 5:1. The samples were then heated to 90°C for 2 min and placed on ice until ready to load. A standard 4.5% (w/v) acrylamide (19:1 acrylamide:bis acrylamide) gel was used for all reactions. Occasionally, primers and

templates were posted to the Protein and Nucleic Acid Chemistry Laboratory (PNACL) at the University of Leicester and the reactions performed on site. Again the results were posted electronically in the form of ABI electropherograms.

2.13 Analysis of sequence

2.13.1 ABI Prism™Edit View

ABI Prism™ Edit View was used to check the parameters used for the gel run, to allow the signal strengths to be evaluated and also to allow the quality of the sequence data to be verified manually.

2.13.2 Seqman II

Contiguous sequences were generated using Seqman II (DNASTAR™, UK). The ABI electropherograms were loaded into the program, filtered to remove any low quality data and then assembled to generate a consensus sequence. Finally any anomalies were rectified by evaluation of the trace data.

2.13.3 Genejockey II

Genejockey II sequence processor (Biosoft, Cambridge, UK) was used for a wide variety of purposes. For example, nucleotide sequences (not trace data) were aligned, translated or restriction sites identified. Peptide sequences were also analysed as appropriate. For example, peptide plots could be derived and amino acid sequences compared.

2.13.4 Internet tools

2.13.4.1 National Center for Biotechnology Information (NCBI)

Query sequences were routinely compared to those at the NCBI (http://www.ncbi.nlm.nih.gov/) using the Basic Local Alignment Tool (BLAST) search engines. Sequences of interest were downloaded and manipulated as Genejockey II files.

2.13.4.2 *B. pertussis* genomic databases and tools

Query sequences were compared to the incomplete *B. pertussis* genome database at:

(http://www.sanger.ac.uk/projects/B_pertussis).

Contigs were searched using the Alpha-powered BLAST search engine using the TBLASTN blast executable at:

(http://www.sanger.ac.uk/Projects/B_pertussis/blast_server.shtml).

Contigs of interest were downloaded and manipulated as Genejockey files.

The PEDANT web site was also used to identify and further characterise *B. pertussis* genes. A list of predicted open reading frames (ORFs) can be found at:

http://pedant.mips.biochem.mpg.de/cgi-bin/p2/wwwfly.pl?set=Bpertussis&page=index

All predicted membrane protein ORFs were identified using PIR (protein function identification) keywords and selecting the membrane protein hyperlink.

2.14 Protein analysis

Unless otherwise stated, the composition of all reagents and buffers can be found in Appendix I.

2.14.1 Fractionation of bacterial cells

Bacteria were harvested from 500 ml of broth culture at 5,000 x *g* for 15 min in a Sorvall superspeed centrifuge (rotor GS-3). A suspension of pelleted bacteria was made in 50 ml of envelope buffer (10 mM sodium phosphate, pH 7.2). The cells were disrupted using a mechanical cell disrupter (One Shot, Constant Systems, Warwick, UK) set to 15 Kpsi. The resultant lysate was centrifuged at 2500 x *g* for 5 min in a Sorvall superspeed centrifuge (rotor SS34) to pellet large debris, unbroken cells and insoluble material. The supernate was centrifuged in a Sorvall ultracentrifuge (OTD-COMBI) at 100,000 x *g* for 45 min. The resulting supernate contained the soluble content of the cytoplasm and periplasm. The pellet obtained from this centrifugation, which contained cell envelopes, was rinsed in envelope buffer and the tubes dried upside down on tissue paper for 10 min. It was then resuspended in 0.5 ml of envelope buffer containing sodium N-laurlylsarcosinate (sarcosyl) 0.5% (w/v) and then mixed by vortexing every 5 min for 30 min to differentially solubilise the cytoplasmic membrane proteins. A further centrifugation was performed as described previously. The final supernate contained the inner membrane proteins and the final pellet (outer membrane) was resuspended in 0.5 ml envelope buffer containing sarcosyl 0.5% (w/v) which could be solubilised in protein loading dye (section 2.15.8).

2.15 Expression of recombinant proteins

2.15.1 Expression of IPTG-induced recombinant proteins

An overnight culture of *E. coli* strains containing expression constructs (Tables 4 and 6), was diluted 1 in 100 in 2 L of 2X YT or LB containing appropriate antibiotics. The culture was then shaken vigorously at 37°C until an ODλ650nm of 0.5-0.8 was obtained. IPTG was then added to a concentration of 1 mM and the culture again shaken at 37°C for 5 h. Cells were harvested at 5,000 x *g* for 15 min in a Sorvall superspeed centrifuge (rotor GS-3).

2.15.2 Expression of proteins in pBAD/gIII

An overnight culture of *E. coli* strain LMG194 containing pBAD(PCT1) (Table 4), was diluted 1 in 100 into 2 L of minimal medium containing 0.2% (w/v) glucose and 100 µg/ml ampicillin. The culture was shaken vigorously at 37°C to an ODλ650nm of approximately 0.5. L-arabinose was added to the flask to a final concentration of 0.02% (w/v) and the culture returned to 37°C with vigorous shaking for 3 h. Cells were harvested at 5,000 x *g* for 15 min in a Sorvall Superspeed centrifuge (rotor GS-3).

2.15.3 Urea extraction of inclusion bodies

After expression of recombinant protein as described above, the entire culture was centrifuged at 5,000 x *g* for 15 min in a Sorvall Superspeed centrifuge (rotor GS-3) The cell pellet was resuspended in 20 ml of Buffer A (Appendix I). The cell suspension was disrupted using a One-Shot cell disrupter and the resultant lysate was centrifuged at 17,000 x *g* in a Sorvall Superspeed centrifuge (rotor SS34) for 30 min to allow inclusion bodies to be pelleted. Such pellets were resuspended in 10 ml of Buffer B (Appendix I) and inclusion bodies solubilised, with tube rotation, for 24 h or until the suspension

cleared. A final ultracentrifugation step of 100,000 x *g* for 3 h was performed and the supernate, containing urea-soluble protein, was retained and stored at -20°C.

2.15.4 Large-scale affinity purification

An 8ml Omni column (Omnitech, Cambridge, UK) containing 8 ml of Nickel-agarose (Qiagen) was attached to an automated Fast Protein Liquid Chromatography (FPLC) apparatus (Controller LCC501, Pharmacia biotech) and equilibrated with Buffer B at 1 ml/min for 1 h. A supercoil was then used to load 8 ml of urea-soluble fraction (section 2.15.3) onto the column. Buffer B was used to wash non-his-tagged proteins from the column until the spectrophotometer (A_{280nm}) detected no proteins in the flow-through. The column was washed with Buffer C, followed by Buffers D and E (Appendix I). Each buffer was pumped through the column until no further protein was detected on the spectrophotometer. Proteins represented by sharp peaks were collected into separate containers. All fractions were retained and analysed by SDS-PAGE.

2.15.5 Small-scale affinity purification

Purification was performed according to the Ni-NTA spin kit (Qiagen), and the method was adapted from that described in the manufacturer's handbook. Briefly, a Ni-NTA column was equilibrated with 600 µl of Buffer B by centrifugation at 700 x *g* for 2 min and discarding the flow through. To the column, 600 µl of urea-soluble protein was added and centrifugation performed as before. The column was then washed three times with 600 µl of Buffer C. The protein was eluted twice in 200 µl of Buffer E and the eluates combined.

2.15.6 Dialysis

Dialysis tubing with a molecular weight cut-off of 12-14 KDa was stored in ethanol 100% and boiled in distilled water immediately before use. The sample was loaded into the tubing, which was sealed with mediclips and placed into 1000 volumes of dialysis buffer (phosphate-buffered saline [PBS] or 1M urea in PBS). The dialysis was performed overnight with stirring at 4°C.

2.15.7 Protein quantification

The Pierce protein estimation kit (Pierce, USA) provided a colourimetric method for estimating protein concentration, according to the microtiter plate protocol. Briefly, Reagent A and Reagent B were combined at a 50:1 ratio to prepare Working Reagent. Bovine Serum Albumin (BSA) (Sigma) and urea-extracted test proteins were diluted appropriately (range 10-2000 µg/ml) and 10 µl was added to the microtitre plate wells. To each well of a microtiter plate 200 µl of Working Reagent was added and the plate incubated at 37°C for 30 min. The absorbance at 562 nm was determined with a microtiter plate reader and protein concentrations estimated from a standard curve.

2.15.8 Sample preparation for electrophoresis

The sample was made up to a total volume of 50 µl and 10 µl 6x protein loading dye (Appendix I) was added. Samples were stored at -20°C until loading. Immediately before use, samples were boiled in a water bath for 10 min. A 10 KDa protein ladder (Gibco BRL) was used according to the manufacturer's instructions.

2.15.9 SDS-PAGE

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed according to the method of Laemmli (1970) in a vertical electrophoresis gel tank (model SE600, Hoefer Scientific Instruments, California, USA). The composition of all solutions can be found in Appendix I. Glass plates were assembled according to the manufacturer's instructions. Resolving gels, containing 7.5-14% acrylamide, were poured between the assembled plates until the gel was 4cm below the top of the plates and ethanol 100% was used as an overlay. The gel remained undisturbed for approximately 1 h at room temperature until set. The ethanol was poured off and the gel surface washed with distilled water. A stacking gel containing 6% acrylamide was poured onto the polymerised separating gel and a comb was placed into the gel solution. Again the gel was left for approximately 1 h at room temperature and the comb removed. The wells were immediately filled with 1x electrode buffer and assembled into the electrophoresis gel tank. The lower buffer reservoir was stirred with a magnetic stirrer during vertical gel electrophoresis. The gel was run at 20 mA until the gel entered the separating (lower) gel when the current was increased to 30 mA. The gel was removed when the tracker dye reached the bottom of the separating gel (approximately 3 h). The gel was removed from the plates and stained or blotted as appropriate. For staining Coomassie blue stain (Appendix I) was utilised, for a minimum of 2 h on a rotating platform. The stain was then replaced with destain solution until the background was decolourised.

2.16 Western blots

2.16.1 Blotting protocol

After SDS-PAGE, proteins were transferred to immobilising nitrocellulose membrane (Hybond-C, Amersham Pharmacia Biotech) according to the method described by

Towbin *et al.* (1979). Briefly, the stacking gel was removed and the separating gel soaked in electroblotting buffer (Appendix I) for 5 min. The blot was then assembled and placed in electroblotting apparatus (Trans-Blot Cell, Bio-Rad, USA) containing electroblotting buffer with magnetic stirring.

2.16.2 Development of blots

After transfer onto the membrane, proteins were visualised by staining in Ponceau S solution (Sigma) for 5 min. Background was removed by destaining in water for 2 min and the position of the molecular weight markers marked with indelible ink. The membrane was then destained in PBS for 10 min and incubated in blocking buffer (Appendix I) for 1 h at room temperature with gentle agitation. Primary antibody was diluted appropriately in blocking buffer and incubated with the membrane for 1 h at room temperature. The membrane was then washed for 15 min in 2 changes of PBS at room temperature. Horse radish peroxidase-conjugated secondary antibody (anti-mouse or anti-rabbit as appropriate, Scottish Antibody Production Unit, UK) was then diluted 1 in 1000, added to the membrane and incubated for 1 h at room temperature. Again the membrane was washed for 2 x 15 minutes in PBS at room temperature. The membrane was then immersed in freshly prepared 3, 3'diaminobenzidine (DAB) (Sigma) solution for approximately 2 min and the staining reaction stopped by washing in distilled water.

2.17 RNA analysis

2.17.1 Extraction of total RNA from *B. pertussis*

B. pertussis cells were harvested from mid-log phase culture in CL medium by centrifugation at 8,000 x *g* for 15 min in a Sorvall superspeed (rotor GS-3). Guanidine thiocyanate (Sigma), 10 volumes, were added to 1 volume of bacterial pellet and the bacteria resuspended by vortexing. The cells were lysed with a mechanical cell

disrupter. The lysed cells were then added to an equal volume of acid phenol (Tris-buffered, pH 4.5, Sigma) and heated to 60°C for 10 min. The reaction mixture was cooled to room temperature and 0.5 volumes of chloroform was added. The mixture was then shaken vigourously by hand for 2-3 min and centrifuged for 15 min at 12,000 x *g* at 4°C (megafuge, Rotor: HFA 14.2). The aqueous phase was then removed, added to an equal volume of isopropanol and left at room temperature for 10 min. The RNA was pelleted by centrifugation as previously described and washed with ethanol 70% (v/v). The RNA was then resuspended in DEPC-treated water (Appendix I) and heated to 70°C for 10 min to aid solubilisation of the RNA. The RNA was then stored at -70°C until used.

2.17.2 Reverse transcription-PCR

Reverse transciption-PCR (RT-PCR) was performed using the Access Kit (Promega, UK) according to the manufacturers' instructions using a single tube reaction. The following components were assembled in a thin-walled 0.5 ml tube, on ice.

	µl
Nuclease free water	30
AMV/*Tfl* 5x reaction buffer	10
dNTP Mix (10 mM each dNTP)	1
Primer 1 (50 pmol/ml)	1
Primer 2 (50 pmol/ml)	1
25mM $MgSO_4$	2
AMV reverse transcriptase	1
Tfl DNA polymerase (5 u/µl)	1
RNA sample (0.25 mg/µl)	3
Final volume	**50**

The above reaction was thermocycled as follows:

First Strand cDNA Synthesis

1 cycle	48°C	45 min
1 cycle	94°C	2 min

Second Strand cDNA Synthesis and PCR Amplification

40 cycles	94°C	30 s
	55°C	1 min
	68°C	1 min
1 cycle	68°C	7 min

The products were analysed by agarose gel electrophoresis according to section 2.3

2.18 *In vivo* tests

2.18.1 Preparation of antigens

Crude urea extracts or His-purified preparations were dialysed into 1 M urea (all carboxy terminal portions) or PBS (all other recombinant proteins). These test antigens were diluted to 75 µg/ml in PBS containing 1.5 mg/ml alhydrogel. Purified P.69 antigen (Smithkline Beecham, batch PAC 031) was also diluted to 75 µg/ml in PBS/alhydrogel.

2.18.2 Active immunisation

Randomised groups of male CD1 mice (Harlan Olac, Bicester, Oxfordshire, UK) aged 3-4 weeks, were injected subcutaneously, under light halothane anaesthesia, with 15 µg of test antigen per mouse. One group of mice was vaccinated with $1/10^{th}$ human dose of a tricomponent acellular vaccine (P.69, FHA, PT) obtained from Smithkline Beecham (Batch DPSK 315A2). Another group of mice received only alhydrogel in PBS and a further control group was left unvaccinated. For each test group, a second dose, as above, was administered after a two week interval. The weight of the mice was recorded prior to the vaccinations and at intervals afterwards to check for toxicity of the antigen preparation. The group size in each instance was 5 mice.

2.18.3 Mouse intranasal challenge

B. pertussis 18-323 or Taberman was grown as a lawn on BG plates in a humidified box at 37°C for 24 h. The resultant growth was suspended in 1% casamino acids solution (Appendix I) and adjusted to 10 opacity units (approx 2×10^9 cfu/ml) by comparison with the 5^{th} International Reference of Opacity (Perkins *et al.*, 1973). The challenge suspension of $2x10^7$ cfu/ml was prepared by diluting 1 in 100 in 1% casamino acids and a sublethal dose of $1x10^6$ cfu/mouse (50 µl) was instilled intranasally under light halothane anaesthesia at seven days after the second vaccination. One control group was challenged with casamino acids only. The weight of the mice was recorded prior to challenge and at intervals up to seven days post-challenge. The mice were sacrificed using a carbon dioxide chamber and the lungs removed aspetically. The lungs were then homogenised in 10 ml of 1% casamino acids using a sterile homogeniser (Silverson machines, UK). The samples, 0.1 ml of both neat and a 1 in 100 dilution, were cultured on BG agar at 37°C in a humidified box. Colonies were counted after 72 h.

2.18.4 Statistical analysis of lung count data

The number of colony forming units per lung was extrapolated from the colony counts obtained on BG agar plates. Nominal values of 20 and 10^7 were given to counts outside the detection limits <100 and >$3x10^6$ respectively. The log_{10} counts were calculated and analysed using Minitab statistical analysis package (version 12). Analysis included calculation of the means and 95% confidence limits, one- and two-way analysis of variance and Tukey's pairwise comparisons.

2.18.5 Antibody production

Groups of 5 mice were injected twice subcutaneously as detailed in section 2.18.2. The groups of mice which had not been challenged were sacrificed one week after the second vaccination and exsanguinated. The blood was allowed to clot at 37°C for 1 h and was then removed from the side of the collection tube and placed at 4°C overnight. The sera were cleared by centrifugation at 4,000 x *g* for 10 min in a Biofuge (Rotor HFA 14.2) and stored at -20°C in 200 µl aliquots.

3.0 Results

3.1 Discovery of a prn-like sequence

Prior to this project being undertaken, Roberts *et al.* had identified an 840 bp amplicon, designated MR30, obtained when primers were directed to the 3' C-terminal encoding domain of pertactin on *B. pertussis* (strain BBC29) genomic DNA (M. Roberts, personal communication). The MR30 amplicon was then cloned into pGEMT to generate pGemT(30). The cloned MR30 amplicon had previously been found to have a restriction pattern different to that predicted for *prn*. Specifically, MR30 contained an *Eco*RI site which was not predicted for *prn*. pGEMT(30) had been digested with *Eco*RI and a resultant 640 bp fragment was subcloned to produce pCRII(640). Partial manual sequencing had confirmed that the MR30 sequence was different, yet similar, to *prn* (GenBank accession number J04560) and may represent part of a similar gene (Castro, Coote, Parton; unpublished observation). Plasmids pGemT(30) and pCRII(640) were provided at the start of the project and some preliminary manual sequencing data was also available.

3.1.1 Sequencing of MR30

Automated sequencing was performed on pGemT(30) and pCRII(640) using vector directed primers M13FOR and M13REV (Table 5). As the average read length was approximately 300 bp per reaction at this time, both plasmids were sequenced in both directions to allow overlapping sequence to be obtained. The resultant sequences derived from the electropherograms were assembled to generate a contiguous sequence using Seqman II (DNASTAR™). Each sequencing reaction was performed in duplicate and the electropherograms were edited manually prior to assembly. The contig was assembled according to Figure 4. Alignment of the MR30 consensus sequence with the predicted *prn* amplicon using GeneJockey II can be seen in Figure 5. The MR30 amplicon sequence was 56% identical to the prn 3' DNA (when the primer site was disregarded). The sequence obtained for MR30 was compared to those available at

Genbank using BLASTn and at the time (1997) did not correspond with previously published sequences. It was assumed that the MR30 sequence was part of a gene with similarity to, but distinct from, *prn*.

The remainder of the potential open reading frame was then investigated. The MR30 sequence had not been amplified using high fidelity DNA polymerase and this sequence had therefore to be confirmed using unamplified DNA. The sequences that the primers annealed to during the PCR, would be similar, but not necessarily identical, to *prn*.

Figure 4: A schematic overview of the strategy used to obtain the sequence of the MR30 amplicon. The sequence names are given at the left of the arrows in the box. FOR sequences were obtained with the M13FOR and REV sequences were obtained with M13 REV. Each arrow represents the results of one sequencing reaction.

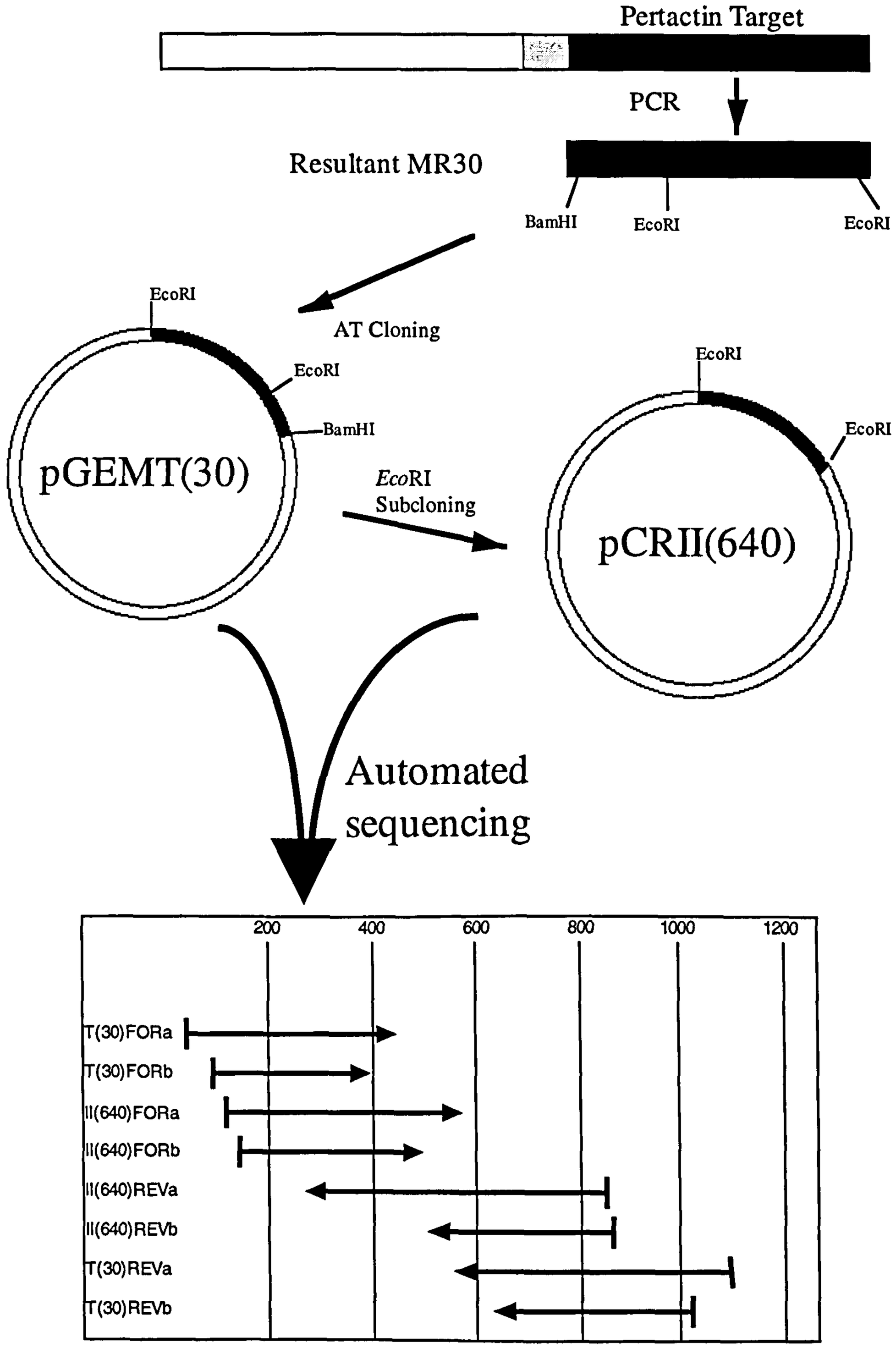

Pertactin Target
PCR
Resultant MR30
BamHI
EcoRI
EcoRI
AT Cloning
EcoRI
EcoRI
BamHI
pGEMT(30)
*Eco*RI
Subcloning
EcoRI
EcoRI
pCRII(640)
Automated
sequencing
200
400
600
800
1000
1200
T(30)FORa
T(30)FORb
II(640)FORa
II(640)FORb
II(640)REVa
II(640)REVb
T(30)REVa
T(30)REVb

Figure 5: Alignment of MR30 Sequence (top) with the corresponding 853 bp of the *Prn* 3' (bottom). Red text denotes sequences obtained from the primers used. The restriction sites *Eco*RI and *Bam*HI included within the primers are underlined. Blue dots highlight nucleotides that are identical in both *prn* and MR30. The *Eco*RI site present in MR30 which initially distinguished it from prn is highlighted in green.

```
          10        20        30        40        50        60        70        80        90       100       110
          |         |         |         |         |         |         |         |         |         |         |
GAATTCGCGTTGTCCAAGCGCCTGGGCGAGTTGCGGCTCGATCCCGGCGCGGGCGGCTTCTGGGGGCGCACGTTCGCCCAG--AAGCAGCAGCTCGACAACAAGGCTGGCCGACGC
      ............................. .. .... . ... ...... ...... ...   ..... ..     ...... . . .. .  .  .  ...
      GCGTTGTCCAAGCGCCTGGGCGAGTTGCGCCTGAATCCGGACGCCGGCGGCGCCTGGGGCCGCGGCTTCGCGCAACGCCAGCAGCTGGACAACCGCGCCGGGCGG--CGC

      120       130       140       150       160       170       180       190       200       210       220       230
      |         |         |         |         |         |         |         |         |         |         |         |
TTCGACCAGAAGGTGTACGGTTTCGAGCTGGGGGCCGACCATGCCATCGCAGGACAGCAAGGGCGCTGGCACGTGGGCGGCCTGCTGGGCTATACCCGCGCAAGGCGCAGCTTCAC
.............  ... .......... ........ ..  . .. .      .. ........ ......  ...   ........ ....     ... .......
TTCGACCAGAAGGTGGCCGGCTTCGAGCTGGGCGCCGACCACGCGGTGGCGGTGGCCGGCGGACGCTGGCACCTGGGCGGGCTGGCCGGCTATACGCGCGGCGACCGCGGCTTCAC

          240       250       260       270       280       290       300       310       320       330       340
          |         |         |         |         |         |         |         |         |         |         |
GATGACGGCGCCGGGCATACCGACAGCGCGCATATCGGGGCCTACGCGGCGTACGTGGCGGACAACGGCTTCTATTTCGATTCGACCCTGCGCGCCAGCCGCTTCGAGAACGACTC
.  ....... ... .. .......... .... ...... ... ..  . ..  . .. .... ... .....  . ..  .... ............... . ..........
GGCGACGGCGGCGGCCACACCGACAGCGTGCATGTCGGGGGCTATGCCACATATATCGCCGACAGCGGTTTCTACCTGGACGCGACGCTGCGCGCCAGCCGCCTGGAGAATGACTC

          350       360       370       380       390       400       410       420       430       440       450       460
          |         |         |         |         |         |         |         |         |         |         |         |
ACGGTAACGGCCACCGACGCCGTTTCCGTACGGGGCAAGTACCGGGCCAATGGGGTAGGCGCCACCTTGGAGGCCGGCAAACGTTTCACGTTGCACGACGGCTGGTTCGTCGAACT
. ...  ... .. .....      . ..  ............ .. ....... ...... .  . ..... ...   .. .. ..      ............. .... ..
AAGGTGGCGGGCAGCGACGGGTACGCGGTCAAGGGCAAGTACCGCACCCATGGGGTGGGCGCCTCGCTCGAGGCGGGCCGGCGCTTTACCCATGCCGACGGCTGGTTCCTCGAGCG
```

Figure 5 (Continued)

```
    470       480       490       500       510       520       530       540       550       560       570
    |         |         |         |         |         |         |         |         |         |         |
CAGTCCGAGGTGTCGCTGTTCCATGCCAGCGGCGGAACCTACCGTGCCGCGAACAACCTGTCGGTCAAG--GACGAAGGCGGCACCTCCGCCGTGCTGCGCCTGGGCTTGGCGGCC
... ..... .. .. . ....  .   ......            ....  ....    .   .  ............. ...   ...... . .          ..
CAGGCCGAGCTGGCGGTATTCCGGG---CCGGCGGCGGTGCGTACCGCGCGGCCAACGGCCTGCGGGTGCGCGACGAAGGCGGCAGCTC--GGTGCTGGGTCGCCTGGGCCTGGAG

580       590       600       610       620       630       640       650       660       670       680       690
|         |         |         |         |         |         |         |         |         |         |         |
GGGCGACGCATCGACCTGGGCAAGGACCGCGTGATCCAGCCCTATGCCACCCTGAGCTGGCTGCAGGAATTCAAAGGCGTCACGACCGTTCGCACCAACGGGTACGGGCTGCGCAC
.   .     .  .          . . ..  .    ..      .. .  .  .  . ... . .      . .. . ..                .. .    .    .. . .
GTCGGCAAGCGCATCGAACTGGCAGGCGGCAGGCAGGTGCAGCCATACATCAAGGCCAGCGTGCTGCAGGAGTTCGACGGCGCGGGTACGGTACACACCAACGGCATCGCGCACCC

    700       710       720       730       740       750       760       770       780       790       800       810
    |         |         |         |         |         |         |         |         |         |         |         |
GACCTGAGCGGTGGCCGGGCTGAATTGGCGCTGGGCCTGGCCGCCGCGTTGGGGCGCGGCCACCAGCTCTACACTTCGTACGAGTACGCCAAGGGCAACAAGCTGACCTTGCCTTG
  .       . ...  .   .       .                .....  .  . .. ..  ...   .    . . .  .  .  .         .  .  .          .
ACCGAACTGCGCGGCACGCGCGCCGAACTGGGCCTGGGCATGGCCGCCGCGCTGGGCCGCGGCCACAGCCTGTATGCCTCGTACGAGTACTCCAAGGGCCCGAAGCTGGCCATGCG

      820       830       840       850
      |         |         |         |
ACGTTCCACCT---GGGCTACCGGTACAGCTGGTAAGGATCC
  .  .  ..    ........................
TGGACCTTCCACGCGGGCTACCGGTACAGCTGGTAA
```

3.1.2 Screening a cosmid library using P640

The pCRII(640) subclone was cut using *Eco*RI to generate a 640 bp fragment which was then used as a template for production of a digoxigenin-labelled Southern blot probe, P640, to enable screening of a *B. pertussis* (Taberman) genomic cosmid library prepared previously (Brownlie *et al.* 1986). Cosmid DNA was extracted from each of the five hundred clones in the library by pooling 5 overnight cultures (1 ml of LB containing 10 mg/ml tetracycline) prior to cosmid extraction. Dot blots were performed at high stringency (68°C) with the P640 probe. Cosmids from the pools that hybridised to P640 were then analysed individually. Subsequent Southern blots revealed many false positive hybridisations. Ultimately one clone, labelled cosmid 3, which was positive with P640 was selected for further analysis.

3.1.3 Subcloning of cosmid 3

To obtain large amounts of cosmid 3 for further analysis, plasmid DNA was prepared from 200 ml of an overnight culture of the *E. coli* clone containing cosmid 3. Cosmid 3 DNA was then cut with several different restriction enzymes and the fragments probed in Southern blots using probe P640 to identify positive fragments. These fragments were then subcloned to provide suitable DNA to allow confirmation of the MR30 sequence and flanking sequences. The restriction enzyme pattern of cosmid 3 can be seen in Figure 6a. A Southern blot of the gel shown in Figure 6a hybridised with probe P640 is seen in Figure 6b. The probe hybridised to bands of 5.2 Kb and 9.3 Kb following incubation of cosmid 3 with *Not*I and *Mlu*I.

Figure 6a: Restriction analysis of cosmid 3.

Kb:	1Kb ladder
Lane 1:	Digest of cosmid 3 with *Not*I
Lane 2:	Digest of cosmid 3 with *Apal*I
Lane 3:	Digest of cosmid 3 with *Xho*I
Lane 4:	Digest of cosmid 3 with *Mlu*I
Lane 5:	Cosmid 3, not digested

Figure 6b: Southern blot of digests of cosmid 3 (figure 6a) hybridised with probe P640.

Figure 6a

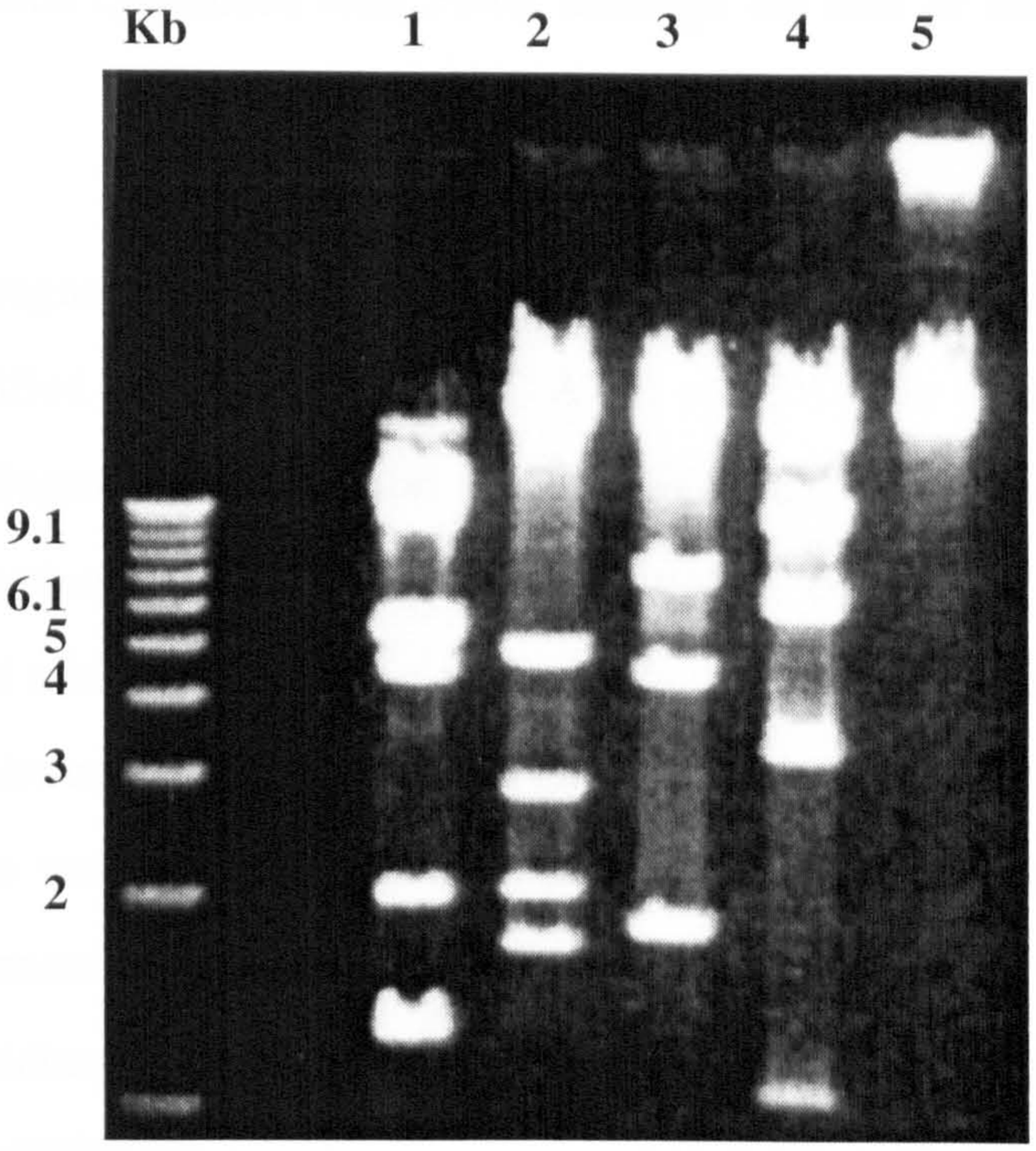
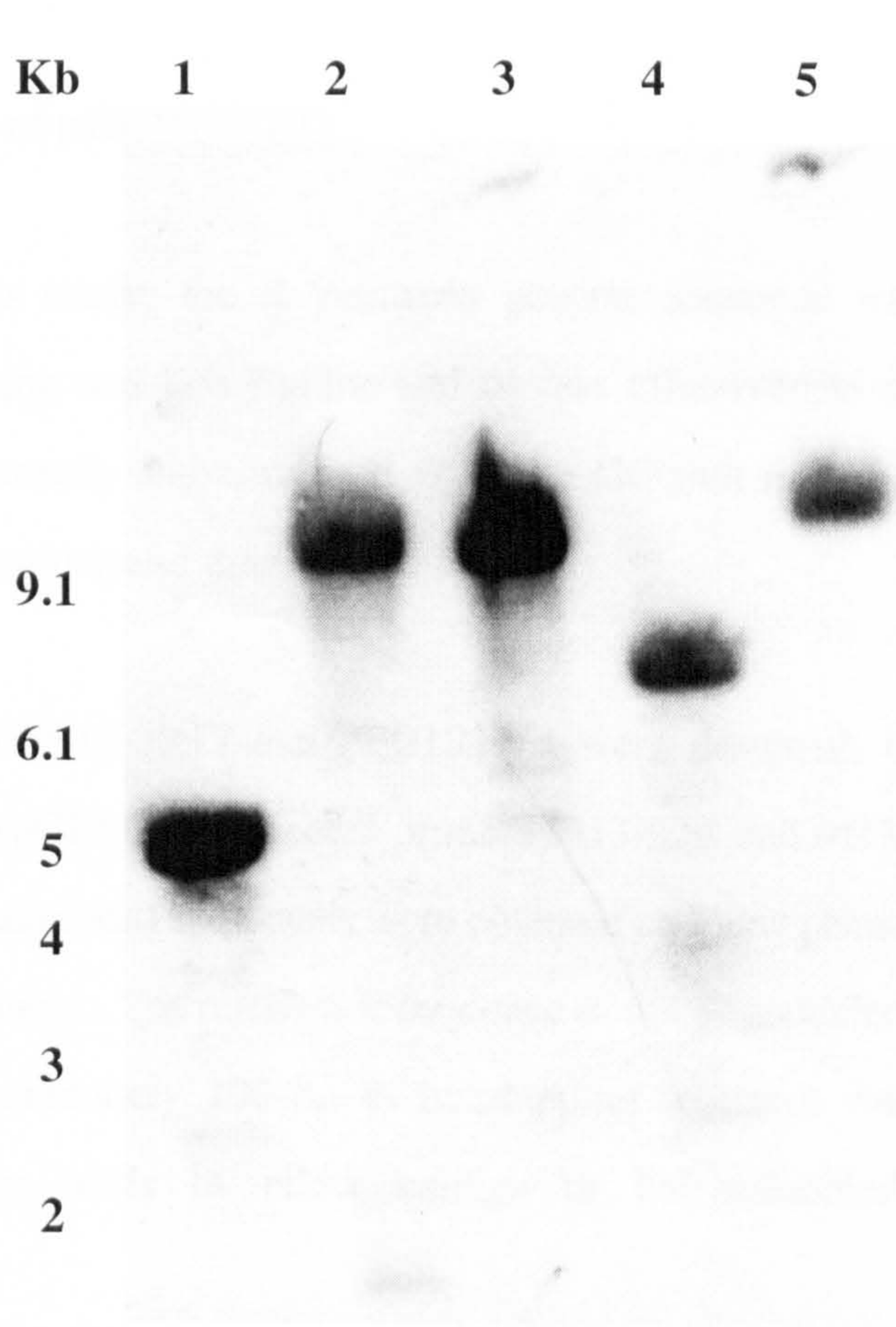


Figure 6b



respectively. The lanes containing DNA treated with *Apa*I and *Xho*I restrictions also contained bands which hybridised with P640, although the size of these fragments was not determined.

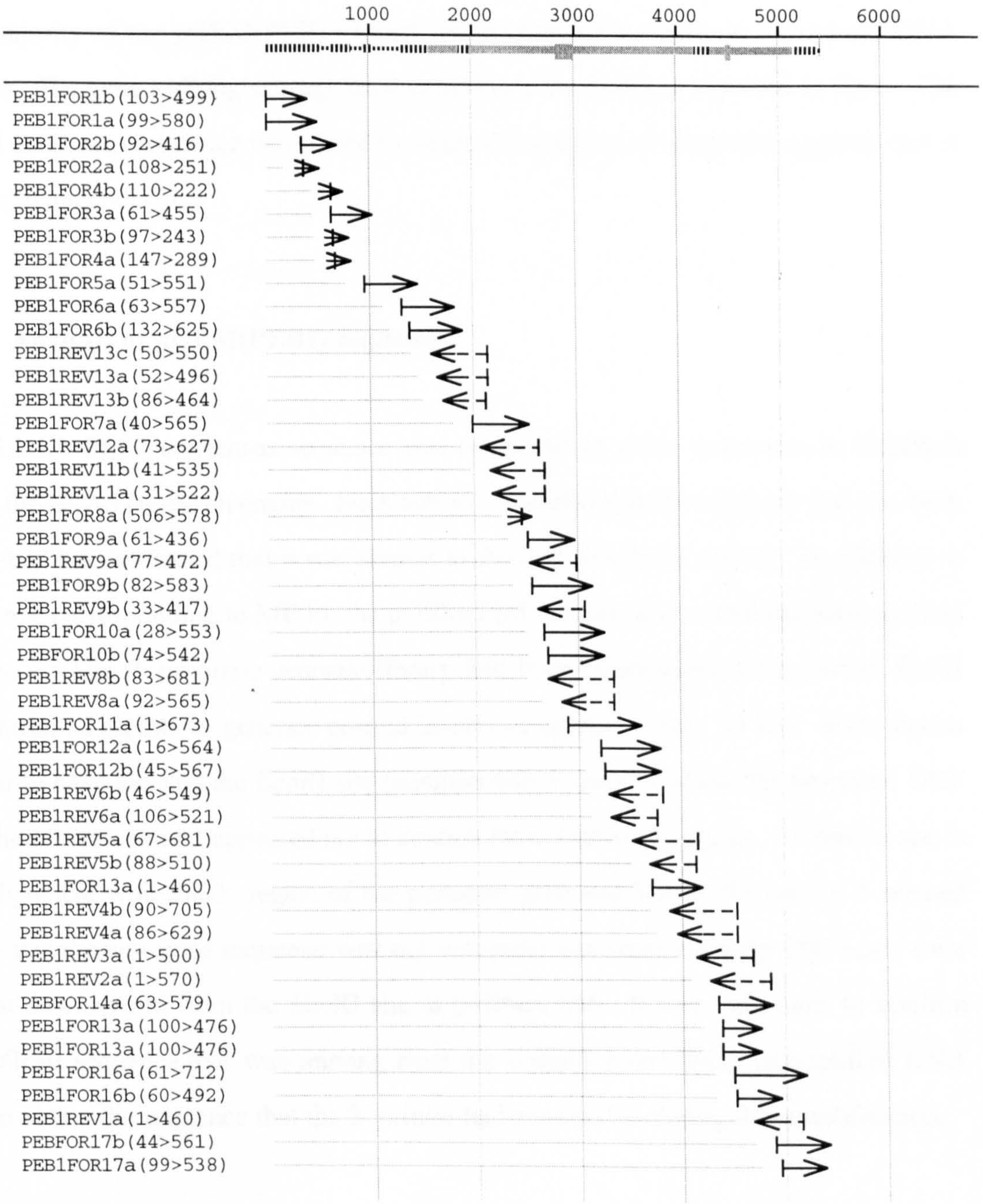
The 5.2 Kb *Not*I fragment that was positive by Southern blot (lane 1, Figure 6a) was excised and gel purified and named "PEB1". A 1:1 and 3:1 ratio of insert : vector was used to ligate the fragment into the *Not*I site of pGEM5Zf using the standard ligation protocol. The ligation mixes were introduced into *E. coli* strain JM109 cells by electroporation and a blue-white screen performed. The 3:1 ratio resulted in 55 blue colonies and 11 white colonies. The 1:1 ratio gave 80 blue colonies and 4 white colonies. Ten white colonies were selected and their plasmids extracted. The plasmids were checked for the presence of an insert by redigestion with *Not*I. The PEB1 insert was confirmed by hybridisation to P640. One clone was then selected for further analysis and named "pGEM(PEB1)".

3.1.4 Sequencing of pGEM(PEB1)

At the time of this study, the *B. pertussis* genome sequence was not available and automated sequencing was less routine and of less effectiveness compared to modern technology. As a result, and combined with the GC-rich nature of the template, the sequencing was laborious and time consuming.

Initially, primers PEB1FOR17 and PEB1REV4 were designed, based on the MR30 sequence. Vector (pGEM5Zf)-directed primers M13FOR and M13REV were also used for sequencing. Automated sequences were obtained and new primers were designed for a gene walking strategy. The nucleotide sequence of the oligonucleotides used are shown in Table 5. Approximately 100 bp of overlapping sequence was generated between individual sequence reads to allow contigs to be assembled using Seqman II.

Figure 7: A overview of the gene walking strategy used to generate the consensus sequence of PEB1. The names of the primers are listed at the left of each sequencing reaction. The numbers in the brackets are the range of sequence which was of medium quality according to Seqman II. Each arrow represents a sequencing reaction, the solid arrows are forward reactions and the dashed arrows are reverse reactions.


1000
2000
3000
4000
5000
6000
PEB1FOR1b(103>499)
PEB1FOR1a(99>580)
PEB1FOR2b(92>416)
PEB1FOR2a(108>251)
PEB1FOR4b(110>222)
PEB1FOR3a(61>455)
PEB1FOR3b(97>243)
PEB1FOR4a(147>289)
PEB1FOR5a(51>551)
PEB1FOR6a(63>557)
PEB1FOR6b(132>625)
PEB1REV13c(50>550)
PEB1REV13a(52>496)
PEB1REV13b(86>464)
PEB1FOR7a(40>565)
PEB1REV12a(73>627)
PEB1REV11b(41>535)
PEB1REV11a(31>522)
PEB1FOR8a(506>578)
PEB1FOR9a(61>436)
PEB1REV9a(77>472)
PEB1FOR9b(82>583)
PEB1REV9b(33>417)
PEB1FOR10a(28>553)
PEBFOR10b(74>542)
PEB1REV8b(83>681)
PEB1REV8a(92>565)
PEB1FOR11a(1>673)
PEB1FOR12a(16>564)
PEB1FOR12b(45>567)
PEB1REV6b(46>549)
PEB1REV6a(106>521)
PEB1REV5a(67>681)
PEB1REV5b(88>510)
PEB1FOR13a(1>460)
PEB1REV4b(90>705)
PEB1REV4a(86>629)
PEB1REV3a(1>500)
PEB1REV2a(1>570)
PEBFOR14a(63>579)
PEB1FOR13a(100>476)
PEB1FOR13a(100>476)
PEB1FOR16a(61>712)
PEB1FOR16b(60>492)
PEB1REV1a(1>460)
PEBFOR17b(44>561)
PEB1FOR17a(99>538)

Problems were encountered with some of the sequencing reactions which resulted in very low sequence quality, which was probably due to the very high GC content of the template DNA. Such difficulties were overcome either by redesigning the primers, adding DMSO (final concentration 5% v/v) or increasing the denaturing time to 40 s. The majority of the pGEM(PEB1) insert was sequenced in duplicate using both DNA strands. The gene walking strategy used to generate the contig is depicted in fig 7. The edited consensus sequence was edited by manual inspection of electropherograms and is shown in appendix II.

3.1.5 Analysis of pGEM(PEB1) sequence

The 5.2 Kb PEB1 consensus sequence was compared to other sequences in GenBank using the BLASTn search engine. Such analysis revealed that the sequence had not been published but confirmed that it was similar to those of *prn*, *brkA* and *tcf*. In addition to sequence corresponding to MR30, the pGEM(PEB1) insert also contained some cosmid sequence. The *B. pertussis* genomic library had been constructed using partial *Eco*RI chromosomal digests to generate cosmid inserts of approximately 50 Kb. Such digests appear to have cleaved the *EcoR*I site (position 646, Figure 5) within the proposed ORF and therefore cosmid 3 appeared not to contain the complete 3' region. As can be see in the alignment with the 3' region of the pertactin gene and MR30 (Figure 5), it seemed likely that the new gene sequence was not complete and approximately 200 bases were absent downstream from the *Eco*RI site at position 646. It was important to confirm the MR30 sequence that was missing from the cosmid insert using unamplified DNA and to obtain the sequence that the 3' primer had annealed to during this amplification.

3.1.6 Retrieval of the 3' sequence of MR30

3.1.6.1 Rescreening of cosmid library

The *B. pertussis* genome library was again screened, this time with a 200 bp probe (P200) which was prepared from the smaller *Eco*RI fragment of pGEMT(30) (Figure 4). Almost every cosmid pool tested appeared positive when this probe was used. Such background was probably due to the size and GC content of this region. Thus, it was not possible to identify a cosmid that genuinely contained the sequence of interest.

3.1.6.2 pUC-ligated single specific primer-PCR

Chromosomal *B. pertussis* (Taberman) DNA (0.5 µg) was digested with either *Pst*I, *Apa*I, *Xho*I, *Hind*III or *Nde*I and ligated to pUC18 (0.75 µg, predigested with appropriate enzyme) to form five gene libraries to serve as templates for SSP-PCR. Primer combinations M13FOR/PEB1FOR16 and M13REV/PEB1FOR16 (Table 5 and Figure 7) were used with an annealing temperature of 53°C. The *Hind*III library SSP-PCR produced two bands for both primer sets. This suggested that target DNA had been cloned into pUC18 in both directions. These bands were cloned into pCRII and their sequences determined. It appeared that the primer PEB1FOR16 was annealing to two positions within one *Hind*III fragment. Unfortunately, the sequences obtained were not similar to that obtained from MR30 and it was not possible to obtain an amplicon that contained the sequence of interest.

3.1.6.3 GenomeWalker™ SSP-PCR

The GenomeWalker libraries were prepared as represented schematically in Figure 8 using *B. pertussis* strain Taberman genomic DNA. Nested SSP-PCR was performed on the gene libraries DL1-DL5 according to section 2.7.2. Figure 9 shows the results of the

secondary PCR on the libraries. The fragment from the *Pvu*II derived library (DL3) of approximately 850 bp (lane3) was purified and ligated into pCRII-TOPO. This ligation mix was then transformed into *E. coli* strain TOP10 and a blue-white screen performed. The 2.2 Kb fragment from the *Dra*I library (DL2) (lane2) was cloned in the same way. Plasmids were prepared from putative recombinants (white colonies) and analysed by *Eco*RI digestion to confirm the presence of the *Eco*RI restriction site within the 3' moiety of the gene. Figure 10 shows the result of this digestion. Lanes 1, 2, 6 and 8 all have a restriction enzyme pattern consistent with that predicted *i.e.* the insert bands add up to approximately 2.2 Kb or 850 bp as appropriate. Lanes 3, 4, 5 and 7 exhibit unexpected restriction patterns for reasons unknown. The plasmids analysed in lanes 1 and 6, named pCRII(DL2a) and pCRII(DL3b) respectively, were chosen for sequence analysis.

3.1.6.3.1 Sequence analysis of pCRII(DL2a) and pCRII(DL3b)

Primers GSP2 and PEB1FOR17 (Table 5) were used to determine the sequence of the pCRII(DL2a) and pCRII(DL3b) inserts. The consensus sequence derived from both inserts had identity to that obtained from pGEM(PEB1) over 381 nucleotides (bases 1960-2341, Figure 5). In addition to this overlap, a further 383 bases were obtained (position 2342-2725, Figure 5) which provided the 3' sequence absent from the subcloned cosmid. It also confirmed the fidelity of the sequence of the MR30 amplicon obtained by PCR (100% identity) and ascertained the actual sequence that the 3' primer had annealed to during the PCR amplification that produced MR30.

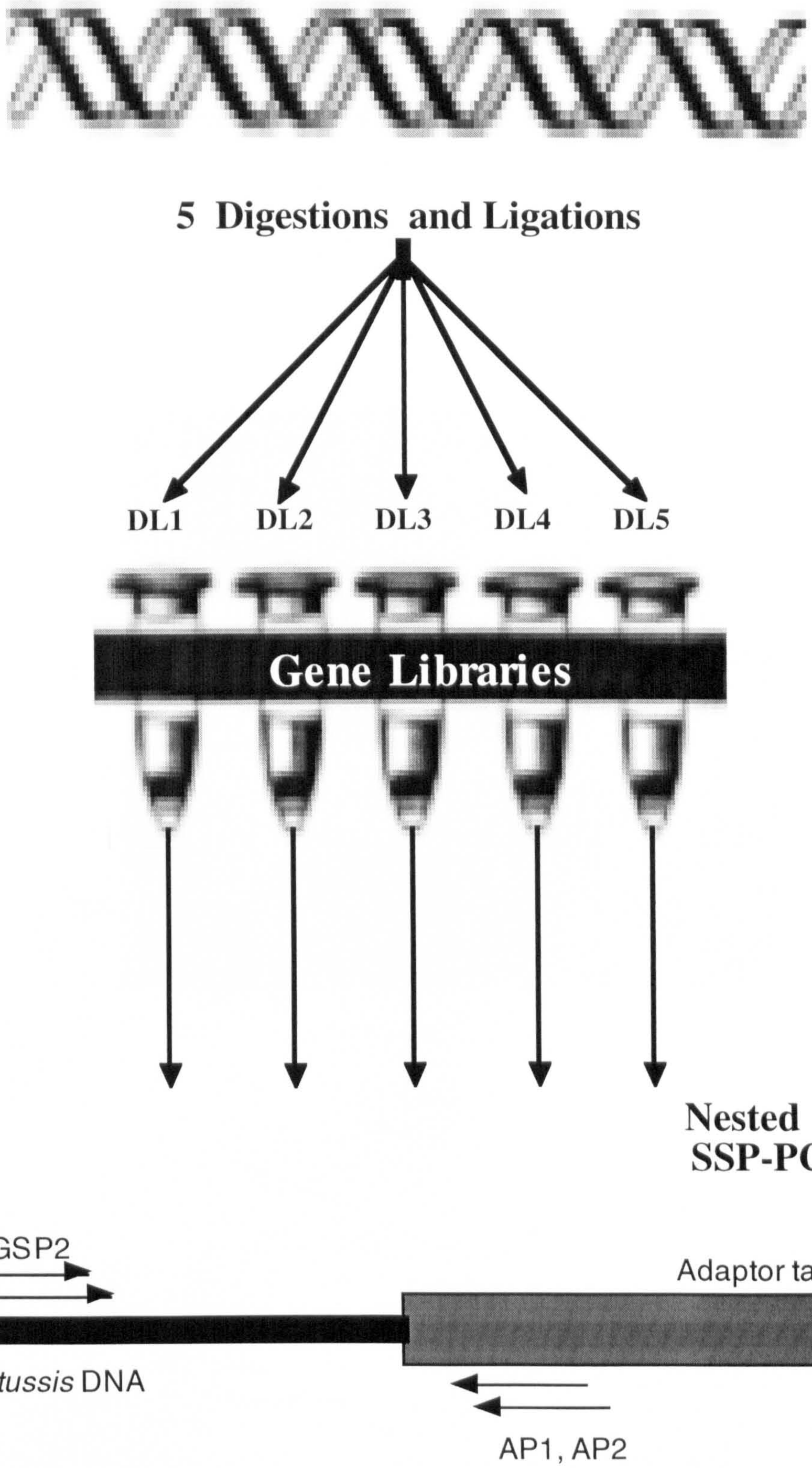


Figure 8: A schematic representation of the GenomeWalker method used to obtain the sequence of the 3' end of MR30. See section 2.8 for a more detailed description of the method.

Figure 9: The amplicons obtained from GenomeWalker™ secondary SSP - PCR

Kb:	1Kb ladder
Lane 1:	Library DL1 (*Eco*RV digest-generated library)
Lane 2:	Library DL2 (*Dra*I digest-generated library)
Lane 3:	Library DL3 (*Pvu*II digest-generated library)
Lane 4:	Library DL4 (*Sca*I digest-generated library)
Lane 5:	Library DL5 (*Stu*I digest-generated library)
Lane 6:	Template-free negative control

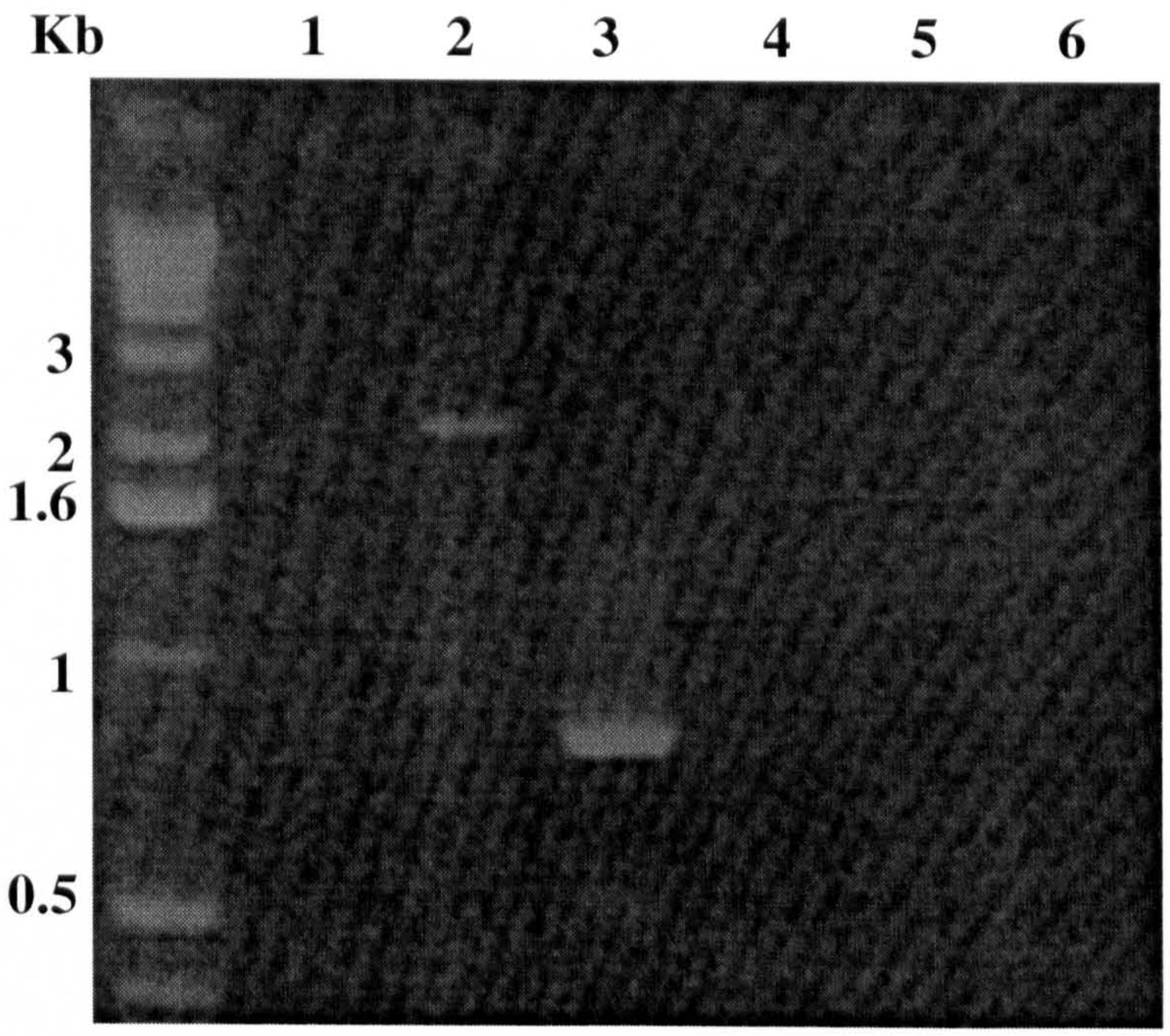

Kb
1
2
3
4
5
6
3
2
1.6
1
0.5

Figure 10: *Eco*RI digests of pCRII constructs containing SSP-PCR amplicons. Note that the vector contains flanking *Eco*RI sites.

Kb: 1Kb ladder

Lane 1: DL1 digest [pCRII(DL2a)]

Lane 2: DL1 digest [pCRII(DL2b)]

Lane 3: DL1 digest [pCRII(DL2c)]

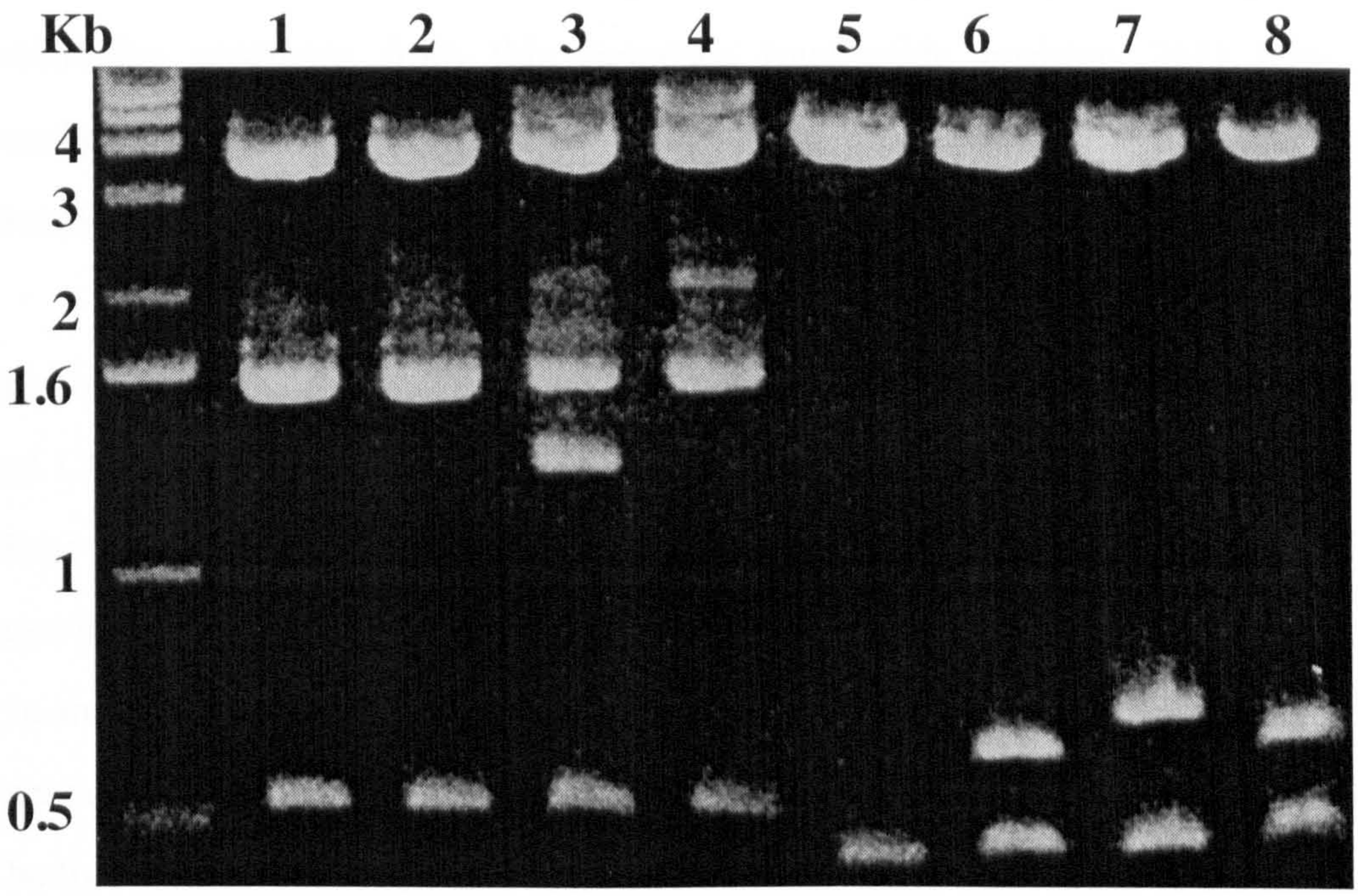
Lane 4: DL1 digest [pCRII(DL2d)]

Lane 5: DL3 digest [pCRII(DL3a)]

Lane 6: DL3 digest [pCRII(DL3b)]

Lane 7: DL3 digest [pCRII(DL3c)]

Lane 8: DL3 digest [pCRII(DL3d)]


Kb
1
2
3
4
5
6
7
8
4
3
2
1.6
1
0.5

3.1.7 Identification of the open reading frame (*bap-5*)

The predicted open reading frame of the pertactin-like sequence is shown in Figure 12. In view of its sequence similarity to *prn* and genes encoding the three other proteins now recognised as belonging to the autotransporter protein family (BrkA, Tcf and Vag8), the sequence was designated *bap-5* (*Bordetella* autotransporter protein 5). The ORF was predicted using amino acid homology with Prn and BrkA, which begins at the serine residue encoded at nucleotide position 300 (Figure 12). This implies that the methionine upstream from this sequence (nucleotide position 261) may be the translational start. The presence of upstream stop codons in this reading frame (positions 43, 103 and 114) also infers that the open reading frame begins at this point.

The sequence obtained was examined for open reading frames by using the TESTCODE and CODONPREFERENCE programmes of the Genetics Computer Group (1991). These programmes predicted that the first methionine (position 261, Figure 12) is the most likely translational start codon. The results of this analysis can be seen in Figures 11a and 11b. However a more recent software package, GLIMMER, suggested a GTG start codon may be the translational start (position 168, Figure 12)(results not shown). The true start however could be identified by N-terminal sequencing of the native Bap-5 protein. However, this may be difficult if Bap-5 encodes a signal sequence which is readily cleaved. This could be determined by cloning and expressing different fragments of *bap-5*.

There are, however, alternative potential start sites (Figure 12) and it would require extensive investigation to ascertain the true ATG translational start codon. There are no translational stop codons within this open reading frame until position 2538. This is at a position consistent with previously characterised *B. pertussis* autotransporters as it is immediately preceded by a consensus outer-membrane localisation motif. A potential transcriptional terminator was identified in the form of mRNA secondary hairpin structure which formed downstream of the predicted ORF (Figure 13).

Figure 11a: The output from CODONPREFERENCE. The nucleotide sequence shown in Figure 12 was analysed. The second reading frame, which begins at position 261, is not interupted by stop codons and the third position GC bias increases significantly in this reading frame. A *B. pertussis* codon frequency table was used for this analysis.

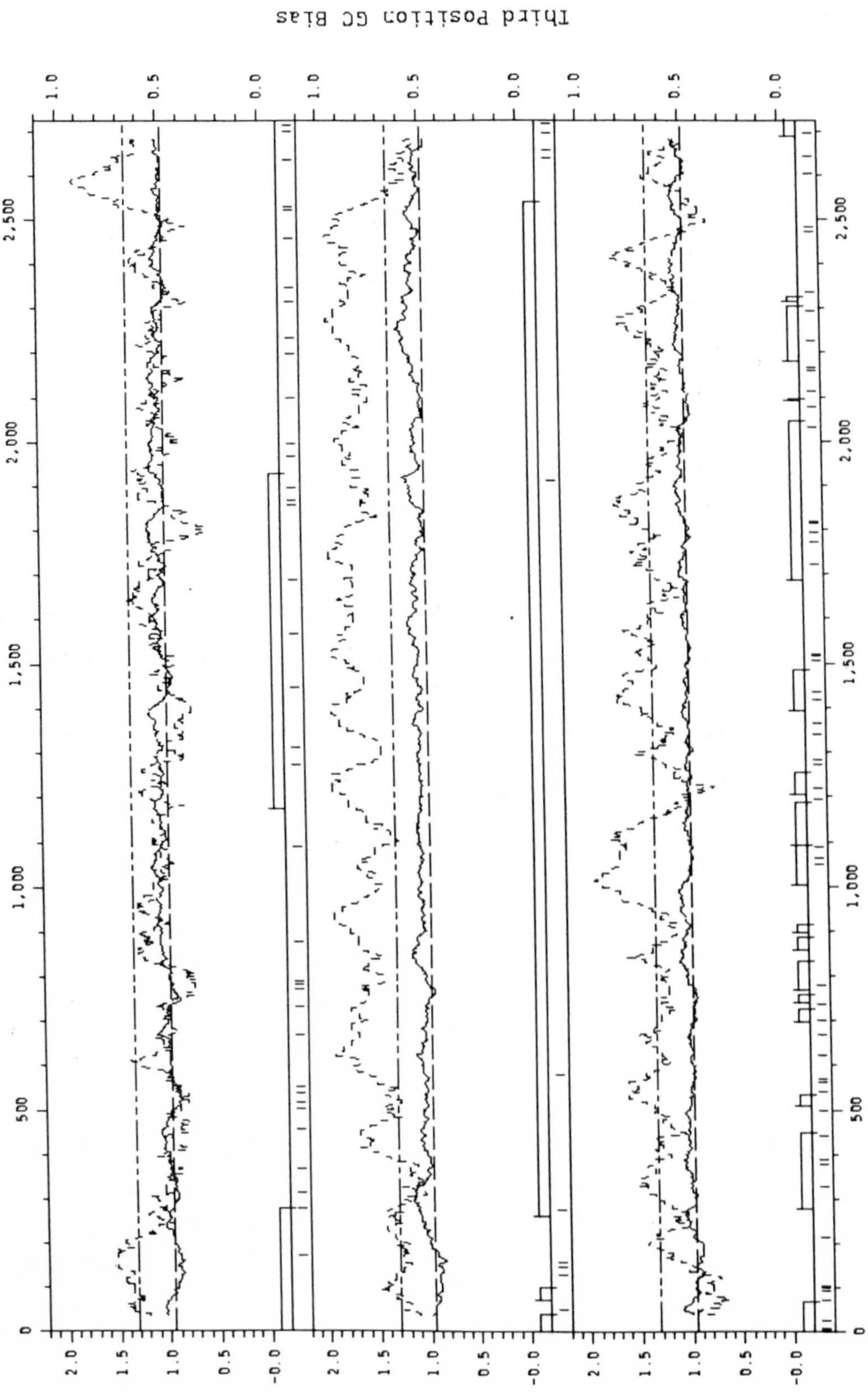

Third Position GC Bias
Codon Preference

Figure 11b: The output from TESTCODE. The nucleotide sequence shown in Figure 12 was analysed. A measure of the non-randomness of the composition of every third base is shown against the nucleotide position. There is a significant rise in the non-randomness of the 3rd position after position 261 (figure 12).

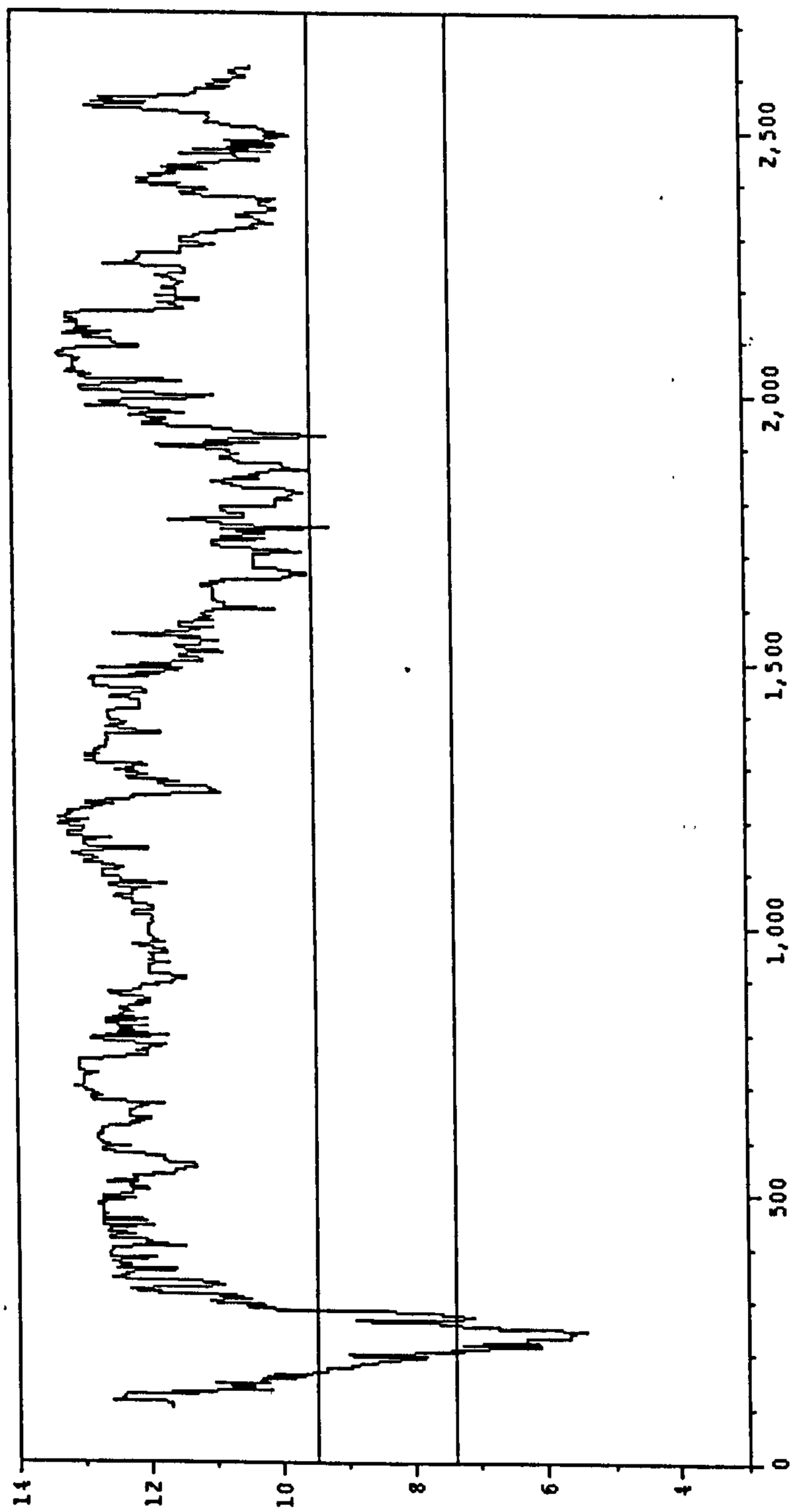

2,500
2,000
1,500
1,000
500
0
14
12
10
8
6
4
TestCode

Figure 12: The nucleotide sequence of *bap-5* and the predicted open reading frame (underlined). The four potential methionine start codons are coloured red. The RGD potential integrin-binding site, and SGSG glycosaminoglycan binding motif are highlighted in yellow boxes. The putative proteolytic processing site and the outer membrane localisation sequence were determined by comparison with Prn, BrkA and Tcf sequences and are also highlighted in yellow. The sequence representing the predicted Rho - independent secondary structure is shown in grey. The Bap-5 specific region, designated NTS, expressed for use in mouse protection tests is coloured green.

```
     10        20        30        40        50        60        70
     |         |         |         |         |         |         |
TCCCGATTATTACGGTGCCGCGGCTGTCTACGCCGGTACGCTGAATATCGAGAATTCCACGGTTCACCATAACTATGC
  P  I  I  T  V  P  R  L  S  T  P  V  R  ·  I  S  R  I  P  R  F  T  I  T  M  R

 80        90        100       110       120       130       140       150
 |         |         |         |         |         |         |         |
GGCCCAGCCGTTCGAAGACGCGGTAGGAGTCGGGGTAACCTCGCTCGGGGATAAGGCCATACTCAACGTTACCGACAG
  P  S  R  S  K  T  R  ·  E  S  G  ·  P  R  S  G  I  R  P  Y  S  T  L  P  T  A

  160       170       180       190       200       210       220       230
  |         |         |         |         |         |         |         |
CGAGGTATCGGGTGCGAGGGGCGCGGTCATCGGTTGGGGGGGGGGGCGAAGCGACATTTACCGATTCGGTCCTGCGTG
  R  Y  R  V  R  G  A  R  S  S  V  G  G  G  G  E  A  T  F  T  D  S  V  L  R  G

     240       250       260       270       280       290       300       310
     |         |         |         |         |         |         |         |
GTTCGGCCTTCGGGCTGTACGCCGAAATGTGCGACACCTGCAGAGATGATGATGGCACCTCGCCTTCGATTCGCGTCC
  S  A  F  G  G  L  Y  A  E  M  C  D  T  C  R  D  D  D  G  T  S  P  S  I  R  V  Q

      320       330       340       350       360       370       380       390
      |         |         |         |         |         |         |         |
AAGGCGGGGTTGTTCAGGGCGGCATGGGTGCAAATAACGTCGCTGTGGTGGCAACAGGGTCTGGAAAGGTCGCGATCG
  G  G  V  V  Q  G  G  M  G  A  N  N  V  A  V  V  A  T  G  S  G  K  V  A  I  E

      400       410       420       430       440       450       460
      |         |         |         |         |         |         |
AGAATGCGGAACTGCTCGGAGCCAGCGGCATGTACGCCACGTTCGGCGCGCAGGTCGATATGAAAGGCGGGCGCATTC
  N  A  E  L  L  G  A  S  G  M  Y  A  T  F  G  A  Q  V  D  M  K  G  G  R  I  L

 470       480       490       500       510       520       530       540
 |         |         |         |         |         |         |         |
TGGCGCACAACACCAATATCCTGGGAAGCCAGGGTTACGCCGATGGTCCCTATGGCGGCGTGGTCGTGACAGAGGACG
  A  H  N  T  N  I  L  G  S  Q  G  Y  A  D  G  P  Y  G  G  V  V  V  T  E  D  G

  550       560       570       580       590       600       610       620
  |         |         |         |         |         |         |         |
GTCAAGTCAACCTGGAGGGCGCCAAGGTCAGTGCAACTGGCCTGGGGGCCGCCGGCTTGTGGTTGCTGGGCGACAAGG
  Q  V  N  L  E  G  A  K  V  S  A  T  G  L  G  A  A  G  L  W  L  L  G  D  K  D

      630       640       650       660       670       680       690       700
      |         |         |         |         |         |         |         |
ACACCAGCCCGCGAGCCAGCCTGCGCAACACCGACGTCCACGGAGAGGTCGCCGCCATTGCGCTGGGGTTCAATGGCG
  T  S  P  R  A  S  L  R  N  T  D  V  H  G  E  V  A  A  I  A  L  G  F  N  G  E

 710       720       730       740       750       760       770       780
 |         |         |         |         |         |         |         |
AGGCGAACATCTCGGGCGGCAGCTTGAGCGTAGAGGATGGGGCCGTGCTCACCACCCTGACGCCCGATGCAGTCGAGT
  A  N  I  S  G  G  S  L  S  V  E  D  G  A  V  L  T  T  L  T  P  D  A  V  E  Y

     790       800       810       820       830       840       850
     |         |         |         |         |         |         |
ATTACTACGACTACGCCTTGTCCATGGAGCATCTGCCAGCTGATGCGCCGTTGACGCCGGTCCGCGTCACGCTGTCCG
  Y  Y  D  Y  A  L  S  M  E  H  L  P  A  D  A  P  L  T  P  V  R  V  T  L  S  D

 860       870       880       890       900       910       920       930
 |         |         |         |         |         |         |         |
ATGGCGCGCGCGCCAGCGGAGAAACGTTGATCGCGCATGGCGGGTTGTTGCCCATGACGCTGCGCTTGAGCAGCGGGG
  G  A  R  A  S  G  E  T  L  I  A  H  G  G  L  L  P  M  T  L  R  L  S  S  G  V
```

```
940        950        960        970        980        990        1000       1010
|          |          |          |          |          |          |          |
TCGACGCCCGCGGCGACATCGTCACGCTGCCGCCTTCCGCGCCGCCCGATTCCGCGGAGCAACCGGATGCCGAGCCGG
  D   A   R   G   D   I   V   T   L   P   P   S   A   P   P   D   S   A   E   Q   P   D   A   E   P   E

1020       1030       1040       1050       1060       1070       1080       1090
|          |          |          |          |          |          |          |
AACCGGATGCCGAGCTGGAACCGGACGCCGCGGCGCAGTCGGACGCCAAGGCGAATGCGCGGGTCATGGCGCAGGTAG
  P   D   A   E   L   E   P   D   A   A   A   Q   S   D   A   K   A   N   A   R   V   M   A   Q   V   D

1100       1110       1120       1130       1140       1150       1160       1170
|          |          |          |          |          |          |          |
ATGGCGGGGAACCTGTTGCCGTGCCGATCCCGGCCCCTTCGCATCCCGATGCCCCGATCGACGTGTTCATCGACAGCG
  G   G   E   P   V   A   V   P   I   P   A   P   S   H   P   D   A   P   I   D   V   F   I   D   S   G

1180       1190       1200       1210       1220       1230       1240
|          |          |          |          |          |          |
GTGCCCAATGGCGGGGCATGACCAAGACCGTCAATGCGTTGCGCATCGAGGACGGCACCTGGACCGTCACCGGGTCGT
  A   Q   W   R   G   M   T   K   T   V   N   A   L   R   I   E   D   G   T   W   T   V   T   G   S   S

1250       1260       1270       1280       1290       1300       1310       1320
|          |          |          |          |          |          |          |
CCACGGTGAACAGCCTGCACCTGCAGGCAGGCAAGGTGGCGTACGCAACGCCTGCCGAAAGCGACGGAGAATTCAAAC
  T   V   N   S   L   H   L   Q   A   G   K   V   A   Y   A   T   P   A   E   S   D   G   E   F   K   H

1330       1340       1350       1360       1370       1380       1390       1400
|          |          |          |          |          |          |          |
ACCTGCGGGTCAAGACCCTCTCGGGAAGCGGCCTGTTCGAGATGAACGCCAGCGCCGACCTGAGCGATGGCGACCTGC
  L   R   V   K   T   L   S   G   S   G   L   F   E   M   N   A   S   A   D   L   S   D   G   D   L   L

1410       1420       1430       1440       1450       1460       1470       1480
|          |          |          |          |          |          |          |
TGGTCGTGTCCGACGAGGCCAGCGGGCAGCACAAGGTGCTGGTGCGAGGAGCCGGCACGGAACCCACCGGTGTGGAAA
  V   V   S   D   E   A   S   G   Q   H   K   V   L   V   R   G   A   G   T   E   P   T   G   V   E   S

1490       1500       1510       1520       1530       1540       1550       1560
|          |          |          |          |          |          |          |
GCCTGACGCTGGTCGAGCTGCCCGAGGGCAGCCAGACGAAGTTCACGCTTGCCAACCGGGGCGGGGTGGTCGACGCCG
  L   T   L   V   E   L   P   E   G   S   Q   T   K   F   T   L   A   N   R   G   G   V   V   D   A   G

1570       1580       1590       1600       1610       1620       1630
|          |          |          |          |          |          |
GCGCGTTCCGCTATCGCCTGACGCCGGACAACGGTGTCTGGGGCCTGGAACGGACCAGCCAGCTTTCGGCCGTCGCCA
  A   F   R   Y   R   L   T   P   D   N   G   V   W   G   L   E   R   T   S   Q   L   S   A   V   A   N

1640       1650       1660       1670       1680       1690       1700       1710
|          |          |          |          |          |          |          |
ACGCGGCCTTGAATACCGGGGGCGTGGGCGCGGCCAGCAGCATCTGGTATGCGGAAGGCAATGCGCTCTCCAAGCGCC
  A   A   L   N   T   G   G   V   G   A   A   S   S   I   W   Y   A   E   G   N   A   L   S   K   R   L

1720       1730       1740       1750       1760       1770       1780       1790
|          |          |          |          |          |          |          |
TGGGCGAGTTGCGGCTCGATCCCGGCGCGGGCGGCTTCTGGGGGCGCACGTTCGCCCAGAAGCAGCAGCTCGACAACA
  G   E   L   R   L   D   P   G   A   G   G   F   W   G   R   T   F   A   Q   K   Q   Q   L   D   N   K

1800       1810       1820       1830       1840       1850       1860       1870
|          |          |          |          |          |          |          |
AGGCTGGCCGACGCTTCGACCAGAAGGTGTACGGTTTCGAGCTGGGGGCCGACCATGCCATCGCAGGACAGCAAGGGC
  A   G   R   R   F   D   Q   K   V   Y   G   F   E   L   G   A   D   H   A   I   A   G   Q   Q   G   R
```

```
      1880       1890       1900       1910       1920       1930       1940       1950
       |          |          |          |          |          |          |          |
GCTGGCACGTGGGCGGCCTGCTGGGCTATACCCGCGCAAGGCGCAGCTTCATCGATGACGGCGCCGGGCATACCGACA
  W  H  V  G  G  L  L  G  Y  T  R  A  R  R  S  F  I  D  D  G  A  G  H  T  D  S

      1960       1970       1980       1990       2000       2010       2020
       |          |          |          |          |          |          |
GCGCGCATATCGGGGCCTACGCGGCGTACGTGGCGGACAACGGCTTCTATTTCGATTCGACCCTGCGCGCCAGCCGCT
  A  H  I  G  A  Y  A  A  Y  V  A  D  N  G  F  Y  F  D  S  T  L  R  A  S  R  F

      2030       2040       2050       2060       2070       2080       2090       2100
       |          |          |          |          |          |          |          |
TCGAGAACGACTTCACGGTAACGGCCACCGACGCCGTTTCCGTACGGGGCAAGTACCGGGCCAATGGGGTAGGCGCCA
  E  N  D  F  T  V  T  A  T  D  A  V  S  V  R  G  K  Y  R  A  N  G  V  G  A  T

      2110       2120       2130       2140       2150       2160       2170       2180
       |          |          |          |          |          |          |          |
CCTTGGAGGCCGGCAAACGTTTCACGTTGCACGACGGCTGGTTCGTCGAACCTCAGTCCGAGGTGTCGCTGTTCCATG
  L  E  A  G  K  R  F  T  L  H  D  G  W  F  V  E  P  Q  S  E  V  S  L  F  H  A

      2190       2200       2210       2220       2230       2240       2250       2260
       |          |          |          |          |          |          |          |
CCAGCGGCGGAACCTACCGTGCCGCGAACAACCTGTCGGTCAAGGACGAAGGCGGCACCTCCGCCGTGCTGCGCCTGG
  S  G  G  T  Y  R  A  A  N  N  L  S  V  K  D  E  G  G  T  S  A  V  L  R  L  G

      2270       2280       2290       2300       2310       2320       2330       2340
       |          |          |          |          |          |          |          |
GCTTGGCGGCCGGGCGACGCATCGACCTGGGCAAGGACCGCGTGATCCAGCCCTATGCCACCCTGAGCTGGCTGCAGG
  L  A  A  G  R  R  I  D  L  G  K  D  R  V  I  Q  P  Y  A  T  L  S  W  L  Q  E

           2350       2360       2370       2380       2390       2400       2410
            |          |          |          |          |          |          |
AATTCAAAGGCGTCACGACCGTTCGCACCAACGGGTACGGGCTGCGCACCGACCTGAGCGGTGGCCGGGCTGAATTGG
  F  K  G  V  T  T  V  R  T  N  G  Y  G  L  R  T  D  L  S  G  G  R  A  E  L  A

      2420       2430       2440       2450       2460       2470       2480       2490
       |          |          |          |          |          |          |          |
CGCTGGGCCTGGCCGCCGCGTTGGGGCGCGGCCACCAGCTCTACACTTCGTACGAGTACGCCAAGGGCAACAAGCTGA
  L  G  L  A  A  A  A  L  G  R  G  H  Q  L  Y  T  S  Y  E  Y  A  K  G  N  K  L  T

      2500       2510       2520       2530       2540       2550       2560       2570
       |          |          |          |          |          |          |          |
CCTTGCCTTGGACGTTCCACCTGGGCTATCGCTACACCTGGTAGCGGCGCATCGCAACAACGGGGTCATCGCGACCCC
  L  P  W  T  F  H  L  G  Y  R  Y  T  W  .  R  R  I  A  T  T  G  S  S  R  P  R

      2580       2590       2600       2610       2620       2630       2640       2650
       |          |          |          |          |          |          |          |
GTTGTTGCATTCCCGGACCGCGCTGCGGTCAGACCAGCCCGGGCCGTCGACCACGCGTGCCCTATAGTAAGGGCGAAT
  C  C  I  P  G  P  R  C  G  Q  T  S  P  G  R  R  P  R  V  P  Y  S  K  G  E  F

      2660       2670       2680       2690       2700       2710       2720
       |          |          |          |          |          |          |
TCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGCCCTATAGTGATAG
  C  R  Y  P  S  H  W  R  P  L  E  H  A  S  R  G  P  N  S  P  Y  S  D
```

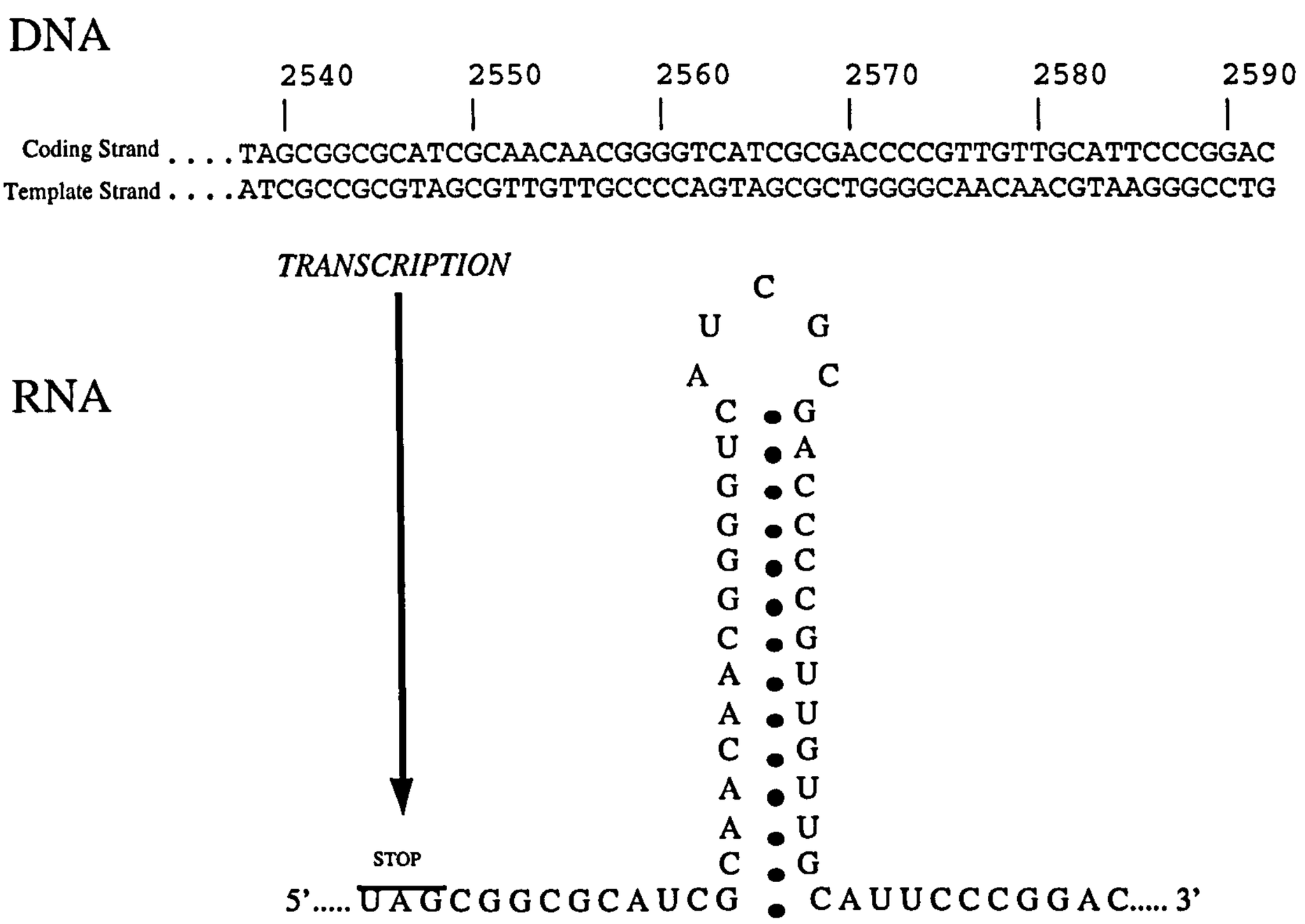


Figure 13: The hairpin structure predicted to occur downstream from the putative Bap-5 open reading frame

The sequence was submitted to GenBank on the 11/08/1998 as a "putative autotransporter protein of *B. pertussis*". The Accession number for this submission is AF081494.

3.1.8 Sequence analysis of Bap-5

As can be seen in Figure 14, the deduced amino acid sequence for the Bap-5 C-terminus is very similar to that described for pertactin and the other, recently identified, *B. pertussis* autotransporters. This alignment strongly suggests that the putative Bap-5 protein should be included in the autotransporter family. Motifs of interest were identified using Genejockey II and are highlighted in Figure 12. These include: A potential integrin binding motif (RGD, position 944-952), a glycosaminoglycan binding site (SGSG position 1347-1358), a proteolytic processing site (ALSKRLGEL, position 1700-1727) and an outer membrane localisation signal (FHLGYRYRW, position 2510-2537). No signal sequence or upstream regulatory elements (*eg.* ribosome binding site or promotor sequences) were identified in the *bap-5* sequence presented in Figure 12. No other potential open reading frames were identified on the PEB1 fragment. The predicted Mw of Bap-5 is 79.5 KDa. If processed at the putative cleavage site, the resultant mature N-terminus is predicted to be 49 KDa and the C-terminus 30 KDa. The N-terminus and C-terminus of Bap-5 were compared with sequences stored in GenBank. In addition to the bordetellae autotransporters, Bap-5 also exhibits similarity to autotransporters in other bacterial genera such as the IgA protease of *N. gonorrhoeae* (results not shown). A more detailed analysis of the similarity between the *B. pertussis* autotransporters can be seen in Figure 15. Bap-5 has highest similarity to the amino and carboxy-terminal domains of Prn. Such similarities with Prn are higher than those seen with the other autotransporter proteins tested. Figure 14 shows the alignment between the *B. pertussis* autotransporter C-terminal domains.

Figure 14: Alignment of the C-terminal amino acid sequences of Prn, BrkA, Tcf and Vag-8 with that predicted for Bap-5. Amino acids coloured red are identical at this position. Amino acids coloured green have similar properties to others at this position. Note that, in many instances, the remaining amino acids in black are identical or similar, at a given position, in at least two members of this family.

```
Bap5 Cterm   AEGNALSKRLGELRLDPGAGGFWGRTFAQKQQLDNKAGRRFDQKVYGFELGADHAIAGQQGR-WHVGGLLGYTRARRSFIDDGAGHT
Prn  Cterm   AESNALSKRLGELRLNPDAGGAWGRGFAQRQQLDNRAGRRFDQKVAGFELGADHAVAVAGGR-WHLGGLAGYTRGDRGFTGDGGGHT
BrkA Cterm   AESNALDKRLGELRLRADAGGPWARTFSERQQISNRHARAYDQTVSGLEIGLDRGWSASGGR-WYAGGLLGYTYADRTYPGDGGGKV
Tcf  Cterm   AEMNALSKRMGELRLTPVAGGVWGRAFGRRQDVDNRVSREFRQTISGFELGADTALPVADGR-WHVGAVAGYTNGRIKFDRGGTGDD
Vag8 Cterm   AEMDVLLRHMSGLHSIGSPGGFWARGLSQRQRLDTGYGPWQKQTVSGIELGLDRRVAGGATTAWSVGMLAGYSETRRDGGAYRAGHV

Bap5 Cterm   DSAHIGAYAAYVADNGFYFDSTLRASRFENDFTVTATDAVSVRGKYRANGVGATLEAGKRFTLHDGWFVEPQSEVSLFHASGGTYRA
Prn  Cterm   DSVHVGGYATYIADSGFYLDATLRASRLENDFKVAGSDGYAVKGKYRTHGVGASLEAGRRFTHADGWFLEPQAELAVFRAGGGAYRA
BrkA Cterm   KGLHVGGYAAYVGDGGYYLDTVLRLGRYDQQYNIAGTDGGRVTADYRTSGAAWSLEGGRRFELPNDWFAEPQAEVMLWRTSGKRYRA
Tcf  Cterm   DSVHVGAYATYIEDGGFYMDGIVRVSRIRHAFKVDDAKGRRVRGQYRGNGVGASLELGKRFTWPGAWYVEPQLEVAAFHAQGADYTA
Vag8 Cterm   HSAHVGAYVSYLNDSGSYVDGVVKYNRFRHGFDIRTTDLKRVDAKHRSHGLGALLRGGRRIDIDGGWYVEPQASVAWFHAGGSRYEA

Bap5 Cterm   ANNLSVKDEGGTSAVLRLGLAAGRRIDLGKDRVIQPYATLSWLQEFKGVTTVRTNGYGLRTDLSGGRAELALGLAAALGRGHQLYTS
Prn  Cterm   ANGLRVRDEGGSSVLGRLGLEVGKRIELAGGRQVQPYIKASVLQEFDGAGTVHTNGIAHRTELRGTRAELGLGMAAALGRGHSLYAS
BrkA Cterm   SNGLRVKVDANTATLGRLGLRFGRRIALAGGNIVQPYARLGWTQEFKSTGDVRTNGIGHAGAGRHGRVELGAGVDAALGKGHNLYAS
Tcf  Cterm   SNGLRIKDDGTNSMLGRLGLHVGRQFDLGDGRVVQPYMKLSWVQEFDGKGTVRTNDIRHKVRLDGGRTELAVGVASQLGKHGSLFGS
Vag8 Cterm   SNGLRVRADGAHSWVLRAGAEAGRQMRLANGNIVEPYARLGWAQELGADNAVYTNGIRHVTRSRGGFAEARVGVGALLGKRHALYAD

Bap5 Cterm   YEYAKGNKLTLPWTFHLGYRYTW
Prn  Cterm   YEYSKGPKLAMPWTFHAGYRYSW
BrkA Cterm   YEYAAGDRINIPWSFHAGYRYSF
Tcf  Cterm   YEYAKGSRQTMPWTFHVGYRYAW
Vag8 Cterm   YEYAKGARFEAPWTLQLGYRYSW
```

Figure 15: A matrix to demonstrate amino acid similarities between the the *B. pertussis* autotransporter domains described to date. The N-termini (N) and C-termini (C) have been compared separately to highlight the extent of the C-terminus homology. The C-terminal and N-terminal domains were chosen on the basis of known or putative cleavage sites. All similarities were determined by BLASTp analysis and are displayed as the % positive (identical and similar) amino acids / extent of the overlap. The extent of the overlap is the window of amino acids which have been analysed. NS represents no significant similarity.

Bap-5: GenBank Acc. No. AF081494
Prn: GenBank Acc. No. AJ006152
BrkA: GenBank Acc. No. U12276
Tcf: GenBank Acc. No. U16754
Vag8: GenBank Acc. No. U90124

	Bap-5	**Prn**	**BrkA**	**Tcf**	**Vag8**
Bap-5	N: 100/480 C: 100/279	N: 45/503 C: 79/279	N: 41/483 C: 62/279	N: NS C: 66/279	N: 40/497 C: 57/265
Prn	N: 45/503 C: 79/279	N: 100/631 C: 100/279	N: 41/636 C: 67/279	N: NS C: 72/279	N: 37/668 C: 58/265
BrkA	N: 41/483 C: 62/279	N: 41/636 C: 67/279	N: 100/731 C: 100/279	N: NS C: 58/279	N: 40/584 C: 56/266
Tcf	N: NS C: 66/279	N: NS C: 72/279	N: NS C: 58/279	N: 100/393 C: 100/279	N: NS C: 55/279
Vag8	N: 40/497 C: 57/265	N: 37/665 C: 58/265	N: 40/584 C: 56/266	N: NS C: 55/279	N: 100/665 C: 100/279

3.1.9 Identification of *bap-5* in *Bordetella* genomes

The digoxigenin-labelled Southern blot probe P640 (section 3.1.2) was used to detect sequences in the genomes of other *Bordetella* species which may be homologous to *bap-5*. Initial problems with high background were overcome by preheating the wash buffers and by reducing the probe concentration. Figure 16 shows the result of the Southern blot analysis of *Bordetella* chromosomal DNA with P640. Strong positive signals were seen in lanes 1-6 which implies hybridisation to the genomes of *B. pertussis* (strains Taberman, BBC29, BBC30 and BP2041), *B. bronchiseptica* and *B. parapertussis*. No signal was detected with *B. avium*. In *B. pertussis* DNA, the size of the positive *Eco*RI fragment was approximately 1.2 Kb. This was smaller than the positive band detected from *Eco*RI digests of *B. bronchiseptica* and *B. parapertussis* (approximately 2.6 Kb). The presence of the same sized bands in the *B. pertussis* strains BBC30 (*prn*$^-$) and BP2041 (*brkA*$^-$) provides evidence that the probe 640 is not hybridising to the *prn* and *brkA* genes.

Figure 16a: Bordetella genomic DNAs digested with *Eco*RI and analysed by agarose gel electrophoresis.

Kb:	1Kb ladder
Lane 1:	*B. pertussis* Taberman (wild type)
Lane 2:	*B. pertussis* BBC29 (wild type)
Lane 3:	*B. pertussis* BBC30 (*prn* deletion of BBC29)
Lane 4:	*B. pertussis* 2041 (*brkA* deletion)
Lane 5:	*B. bronchiseptica* 214 (wild type)
Lane 6:	*B. parapertussis* 5952 (wild type)
Lane 7:	*B. avium* 4480 (wild type)

Figure 16b: Southern blot of bordetella genomic DNAs (figure 16a) with probe P640.

Figure 16a

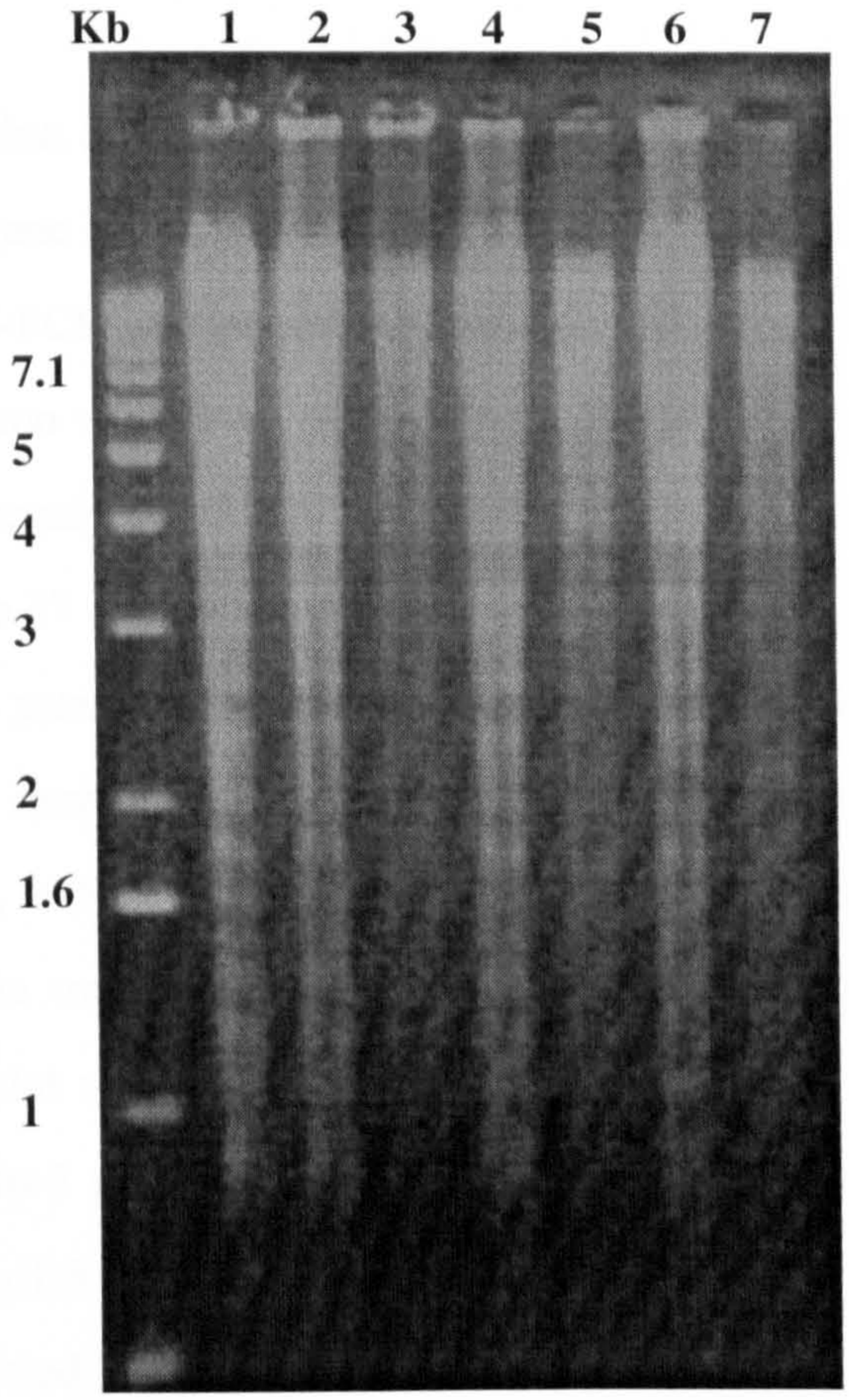
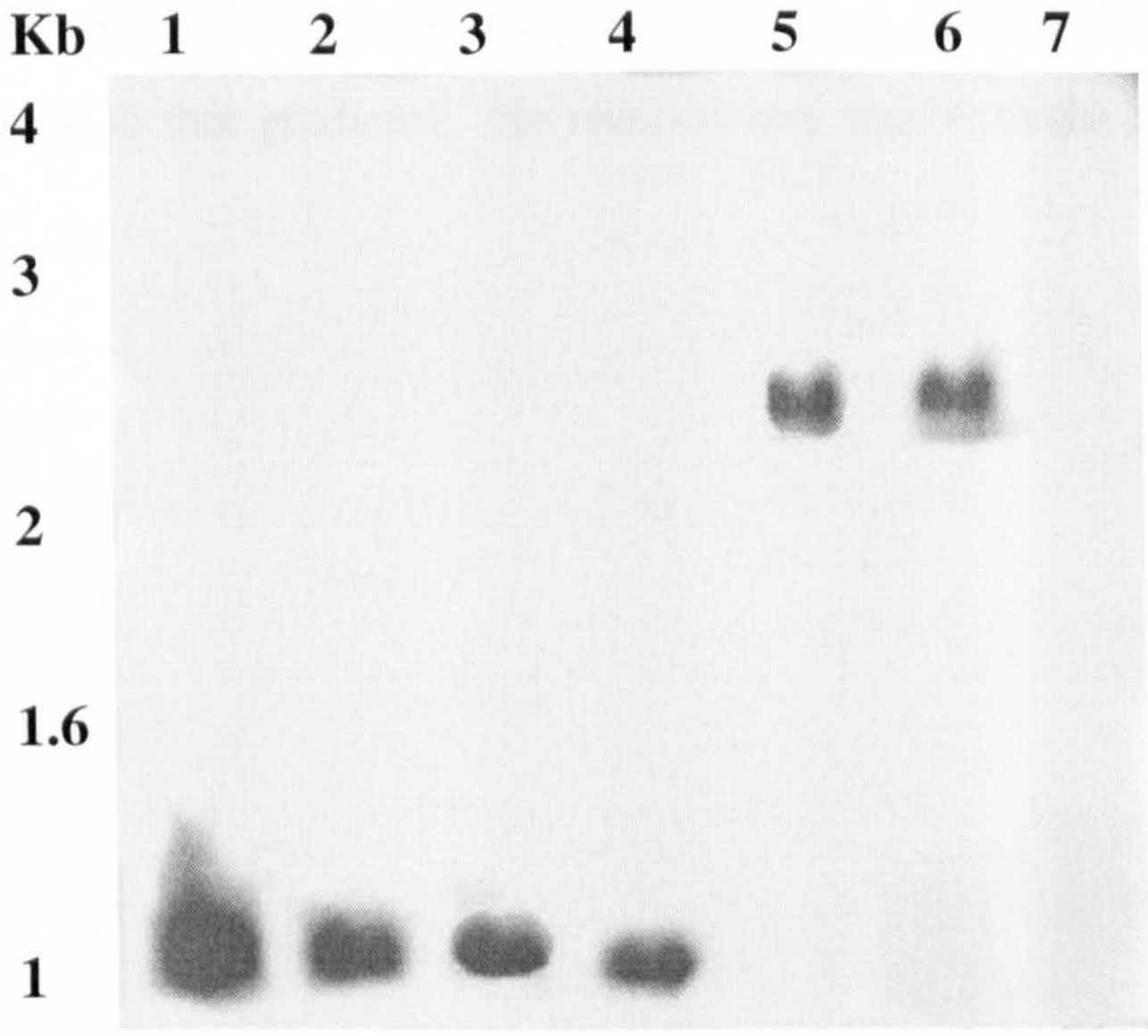


Figure 16b



3.1.10 Expression of *bap-5* in *B. pertussis*

To analyse the expression of *bap-5*, RT-PCR and Western blotting was performed on RNA and outer membrane preparations of *B. pertussis* strains Taberman (wild type) and BP347 (Bvg^-). RT-PCR was performed using the primers NTSFOR and NTSREV (Table 5) and an amplicon was obtained on several occasions. This procedure was not highly reproducible, presumably due to the instability of prokaryotic RNA preparations (section 2.17.1). Figure 17 shows the amplicon obtained from the RNA template from strain Taberman. This product was of predicted size (450 bp) and was not present in the sample with no reverse transcriptase which implies that the template was transcribed from RNA. Also, no amplicon was obtained from the Bvg mutant BP347 which suggests that the transcription of this gene is Bvg-regulated (like Prn, section 1.3.1). The Western blot was performed using pooled mouse antisera which had been raised to the whole Bap-5 protein (sections 3.2 and 2.18.5) and reacted with an outer membrane preparation of *B. pertussis* strain Taberman at 65 KDa, 38 KDa and 30 KDa (Figure 18). The predicted size for unprocessed Bap-5 is 79.5 KDa and the processed forms are predicted to be 49 and 30.5 KDa for the N and C-termini respectively. Therefore, the 30 KDa reaction could represent the C-terminal domain and the 65 KDa could be either the N-terminal portion migrating at higher Mw (as with P.69) or an unprocessed form which is either migrating at lower Mw, or translation has occured from a site different to that predicted. No reaction was seen with the avirulent (Bvg^-) mutant, BP347.

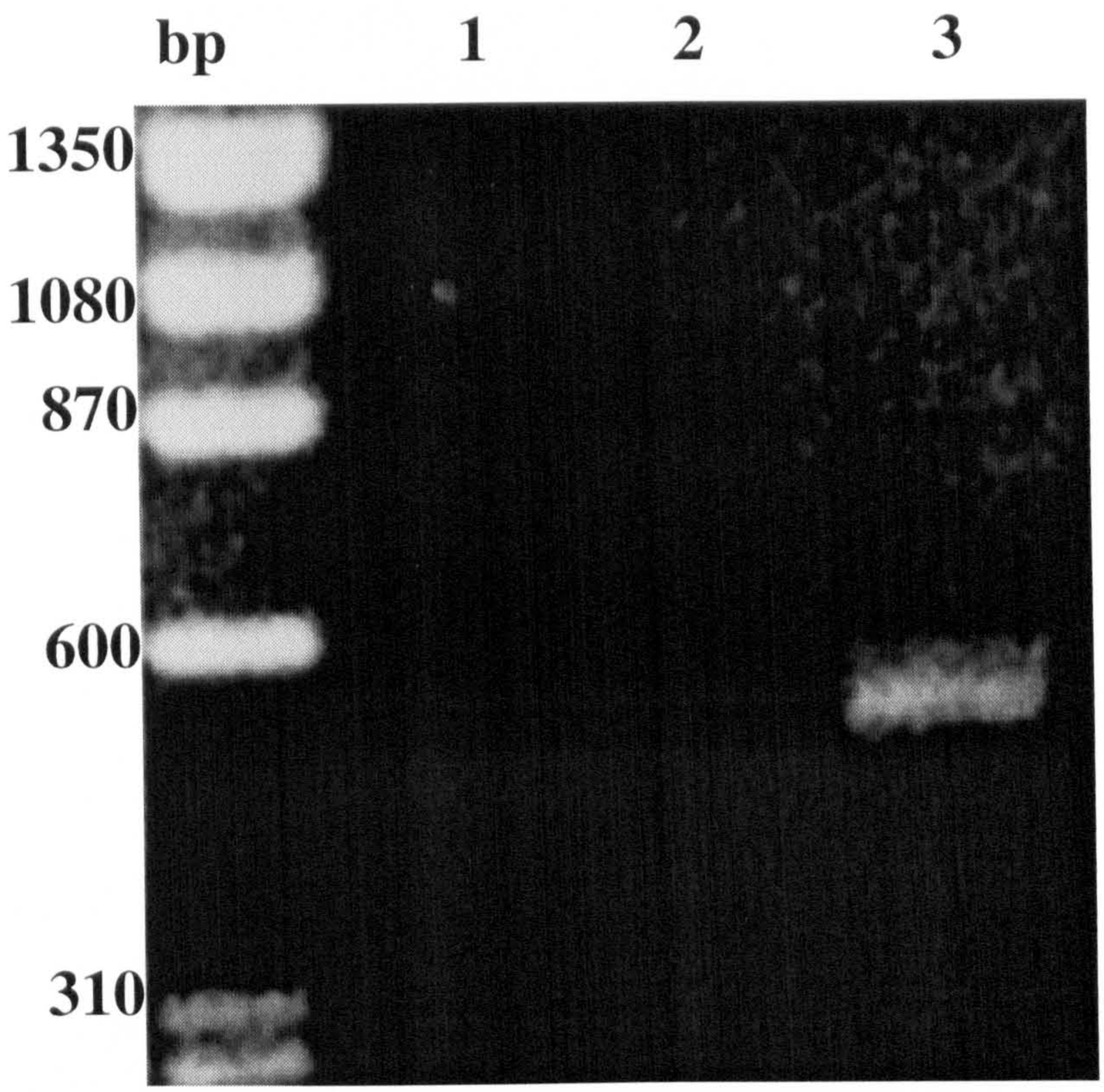
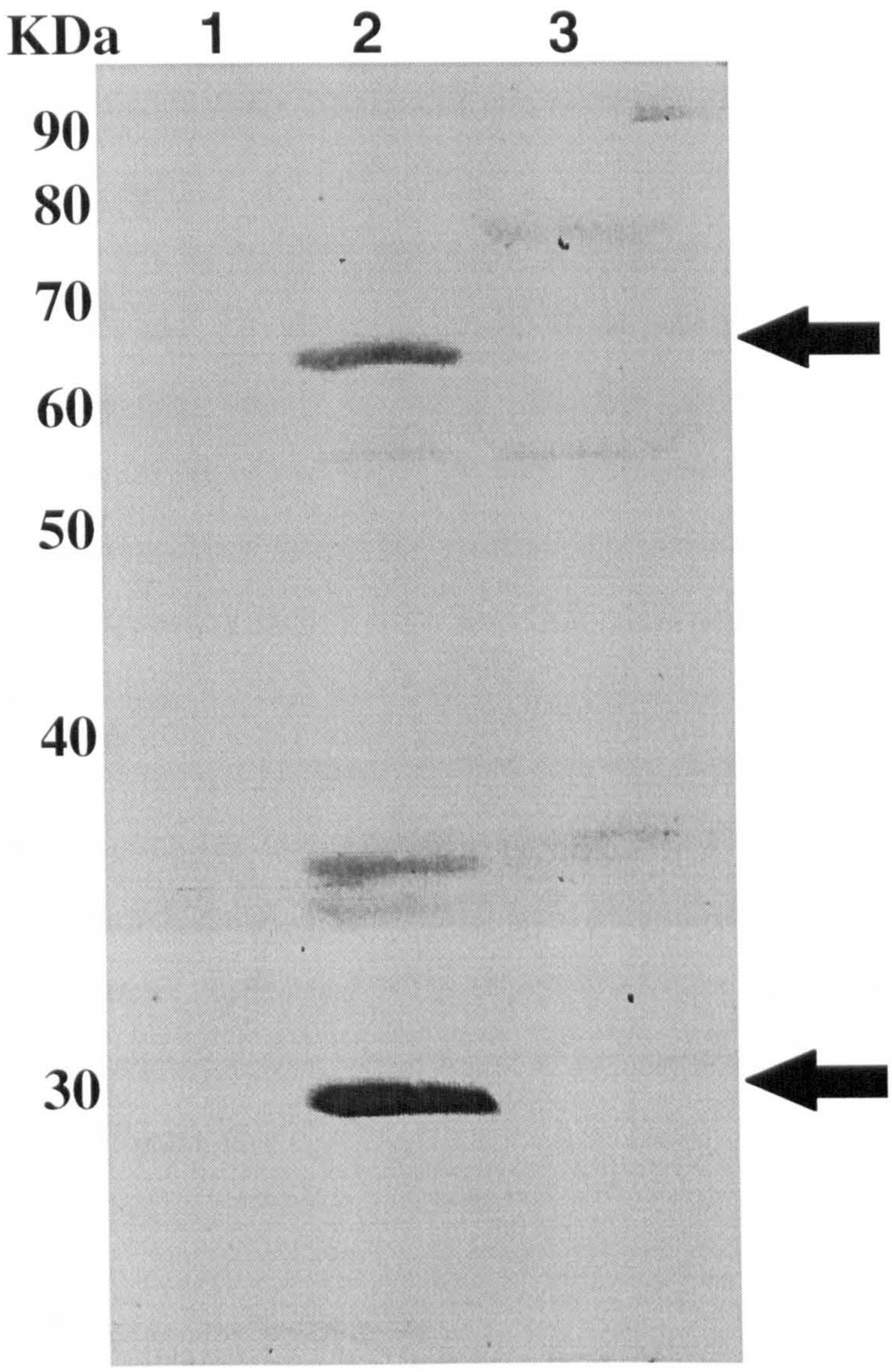

bp
1
2
3
1350
1080
870
600
310

Figure 18: Western blot of outer-membrane preparations using anti-whole Bap-5 sera. Arrows highlight the strongest reactions with the antibody. Non-specific reactions are evident in lane 3 and are of lower intensity than the proposed reactions.

Lane 1: *E. coli* outer membrane preparation (negative control)

Lane 2: *B. pertussis* strain Taberman (wild type) outer-membrane preparation

Lane 3: *B. pertussis* strain BP347 (Bvg^-) outer-membrane preparation


KDa
1
2
3
90
80
70
60
50
40
30

3.2 Recombinant expression of domains from *B. pertussis* autotransporter proteins

The intention was to directionally ligate autotransporter-encoding DNA fragments into pET11a, pET33b (Novagen) or pQE-60 (Qiagen) for high level expression in *E. coli*. Such material would then be purified and used in acellular vaccine formulations for mouse protection tests and for structural analysis, including X-ray crystallography. The true translational start codon of *bap-5* was not known, therefore to allow recombinant expression of the whole Bap-5 protein, constructs were created using each of the four ATG codons at the 5' end of the putative open reading frame (section 3.1.7). Only the construct expressing Bap-5 from the first methionine was used in further experiments. The rationale for this was that all the potential start codons were in the same reading frame and, even if the incorrect start site was chosen, the resultant protein would still contain the complete Bap-5 amino acid sequence. Plasmids expressing the C-terminal portions of Bap-5, BrkA, Prn and Tcf were also created. In addition, a Bap-5 specific N-terminal domain, designated NTS, which bears less similarity with the other known *B. pertussis* autotransporters, was expressed and purified. The NTS sequence is highlighted in green in Figure 12

3.2.1 Creation of expression constructs

All the expression systems were produced according to the schematic overview presented in Figure 19, with the 850 bp C-terminus of pertactin used as an example.

3.2.1.1 Amplification of sequences encoding autotransporter domains

The sequences of all primers used are given in Table 5. For the entire predicted *Bap-5* open reading frame, primers were designed to all 4 potential start codons and to the predicted 3' TGA stop codon (position 2537, Figure 12). Primers were also designed to

amplify the DNA encoding the C-terminal regions of Bap-5, Prn, BrkA and Tcf based on published sequences.

For amplification of DNA encoding the C-terminal domains, an ATG start codon was engineered into the 5' end of the artificial open reading frame to permit translation. Restriction sites were incorporated into the primers to allow directional cloning into the expression vector (Table 6). Care was taken during primer design to ensure that the correct reading frame would be maintained. The template used for each amplification was *B. pertussis* (Taberman) genomic DNA. High fidelity amplification of *bap-5*, from all 4 potential start codons was obtained, (two shown in Figure 20b). In addition, a region of Bap-5, NTS, which had least homology to other described autotransporters was also amplified to allow expression of a Bap-5 specific peptide for antibody production (Figure 20a).

Figure 19: A schematic overview of the methods used to create protein expression constructs.

Figure 19a: PCR amplification of DNA encoding the pertactin C-terminus (850 bp, see arrow) according to Taguchi optimisation methods (Cobb and Clarkson 1994). See Table 5 for primers used. Note that in addition to the major product of predicted size, non-specific amplification products are evident in lane 9 at approximately 1.6 Kb and 2.7 Kb. All reactions were carried out in a 50 µl volume with 5% DMSO, 3U Taq polymerase (Promega), 1x reaction buffer (Promega). The template DNA was from *B. pertussis* strain Taberman. The variables are given in brackets and represent the final reaction conditions.

Kb: 1 Kb ladder

Lane 1: Products from tube 1:

(0.5 pM each primer; 50 ng template; 1 mM $MgCl_2$; 0.1 mM each dNTP)

Lane 2: Products from tube 2

(0.5 pM each primer; 150 ng template; 2 mM $MgCl_2$; 0.2 mM each dNTP)

Lane 3: Products from tube 3

(0.5 pM each primer; 300 ng template; 3 mM $MgCl_2$; 0.3 mM each dNTP)

Lane 4: Products from tube 4

(1 pM each primer; 50 ng template; 2 mM $MgCl_2$; 0.3 mM each dNTP)

Lane 5: Products from tube 5

(1 pM each primer; 150 ng template; 2 mM $MgCl_2$; 0.1 mM each dNTP)

Lane 6: Products from tube 6

(1 pM each primer; 300 ng template; 1 mM $MgCl_2$; 0.2 mM each dNTP)

Lane 7: Products from tube 7

(1.5 pM each primer; 50 ng template; 2 mM $MgCl_2$; 0.3 mM each dNTP)

Lane 8: Products from tube 8

(1.5 pM each primer; 100 ng template; 1 mM $MgCl_2$; 0.3 mM each dNTP)

Lane 9: Products from tube 9

(1.5 pM each primer; 300 ng template; 2 mM $MgCl_2$; 0.1 mM each dNTP)

Lane 10: Template-free control (otherwise equivalent to tube 5)

Figure 19b: Restriction digest of the insert encoding the pertactin C-terminus from pCRII using *Bsp*HI and *Bam*HI. This DNA (at 850 bp, see arrow) was gel purified and ligated into the overexpression vector pET33b (*Nco*I and BamHI sites).

KB: 1 Kb ladder

Lane 1: Restriction of pertactin C-terminus encoding DNA from pCRII

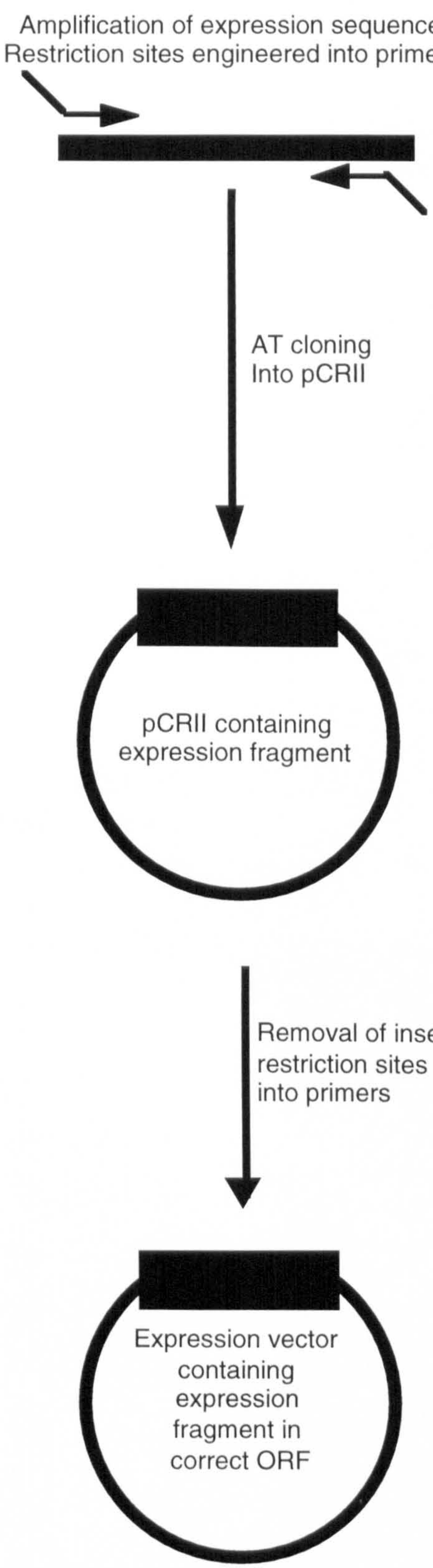


Figure19a

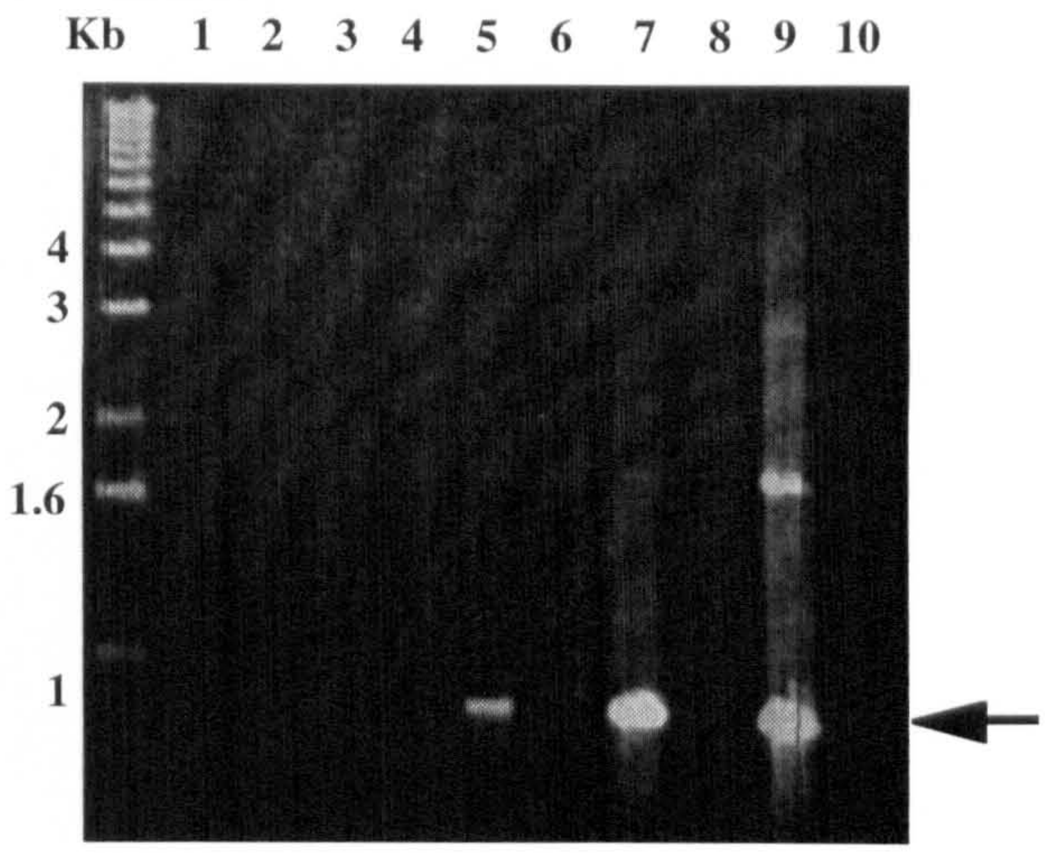


Figure19b

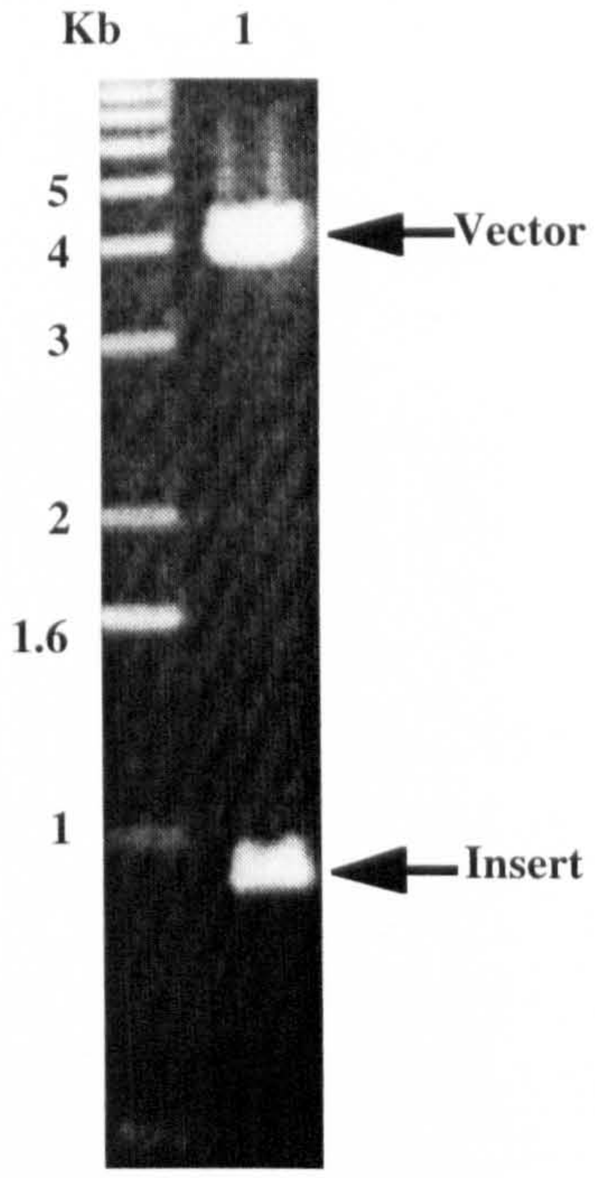


Figure 20a: Amplification of DNA that encodes a specific portion of the Bap-5 N-terminus (NTS)

Kb: 1Kb ladder

Lane 1: Template-free negative control

Lane2: Amplification of NTS-encoding DNA (450 bp)

Figure 20b: Amplification of *bap-5*:

Kb: 1Kb ladder

Lane 1: Amplification of *bap-5* from the 1st ATG (position 261, Figure 12).

Lane 2: Amplification of *bap-5* from the 2nd ATG (position 336, Figure 12)

Lane 3: Template-free negative control with primers designed for amplification from the first ATG of *bap-5*.

Figure 20a

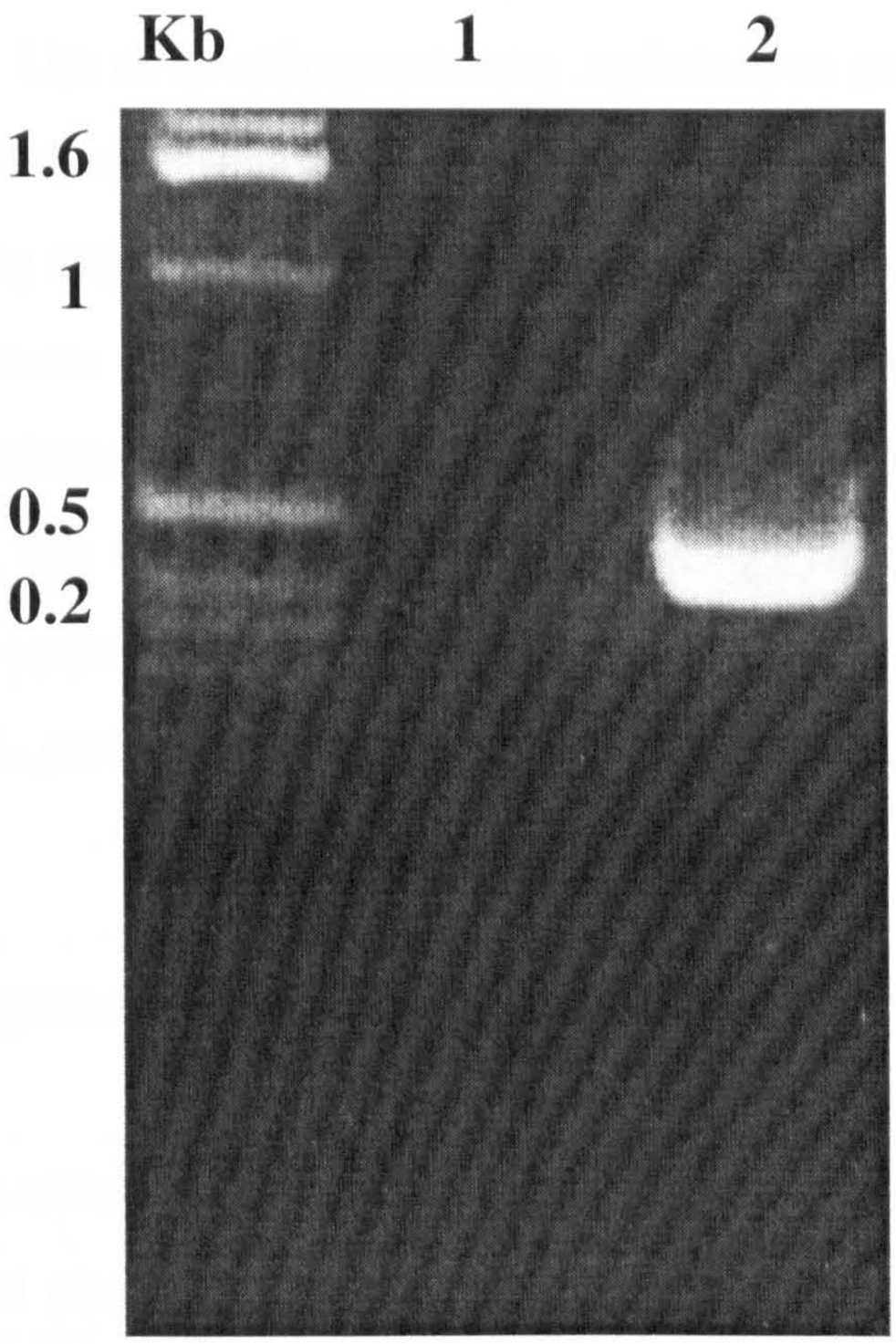
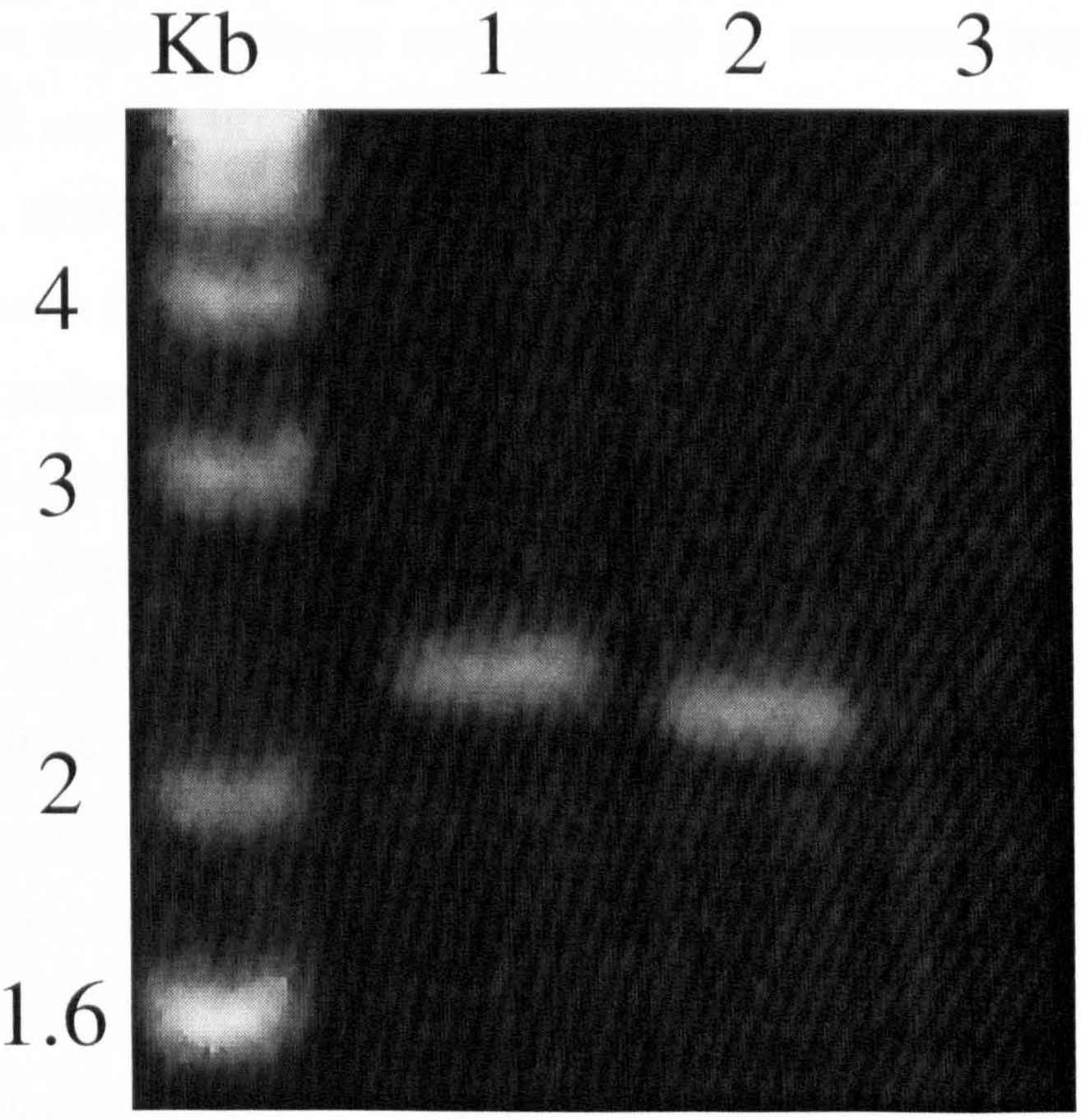


Figure 20b



Attempts to amplify the C-terminus encoding regions of Prn, BrkA and Tcf with proof-reading DNA polymerase were not successful, probably due to the high GC content of the template sequence. Use of an orthogonal array optimisation procedure described by Cobb and Clarkson (1994) enabled optimisation of these reactions and the results of the Prn C-terminus encoding amplification can be seen in Figure 19a. The pattern produced by this array was the same for each C-terminus encoding fragment amplified, *i.e.* with *prn* lanes 5, 7 and 9 produced a band of the expected size (850 bp). The product from lane 5 was chosen on this occasion. The Bap-5 C-terminus-encoding amplicon was amplified later in the project using HotStarTaq™ DNA polymerase (section 2.7.2) which was not available previously.

3.2.1.2 Cloning of expression amplicons into pCRII

To propagate the DNA of interest in a form that could be easily prepared and restricted, the amplicons were gel purified, AT-cloned into pCRII-TOPO and transformed into *E. coli* strain TOP10F'. Plasmids were purified from 10 resultant white colonies and analysed for presence of an insert of correct size by *Eco*RI restriction analysis. The sequences of the cloned amplicons were confirmed by automated sequencing using primers M13FOR and M13REV (Table 5). The *bap-5* amplicons were also sequenced using M13FOR, M13REV, PEB1REV4, PEBREV5 and PEB1FOR12. Fewer primers were needed to cover the whole *bap-5* amplicon than previously required as sequencing reactions were performed at the Protein and Nucleic Acid Chemistry Laboratory (PNACL) at the University of Leicester. Read lengths from PNACL were longer than those previously obtained from the MBSU at the University of Glasgow. However, the sequences were confirmed as identical to those obtained from previous work or those available at GenBank (accession numbers: J04560 [*prn*], U12276 [*brkA*] and U16454 [*tcf*]). From the clones analysed, the fidelity of the non-proofreading Taq polymerase used was excellent. The constructs containing correct insert were named pCRII(PCT1), pCRII(TCF1), pCRII(BRKA1), pCRII(Bp5CT1), pCRII(Bap5M1), pCRII(Bap5M2), pCRII(Bap5M3), pCRII(Bap5M4) and pCRII(NTS) (Table 6).

3.2.1.3 Cloning of expression fragments into overexpression vectors

Initially, pET11a was used to clone the DNA encoding the BrkA C-terminus. However, problems were encountered when trying to ligate other fragments into this vector. It was also realised that a vector encoding an Histidine$_6$ tag (N-terminal for pET33b; C-terminal for pQE-60) to aid purification would be advantageous. The DNA encoding the Prn and BrkA C-terminal fragments were ligated into pET33b for this reason. Again, technical problems with the vector pET33b led to use of another readily available Histidine$_6$ tag-encoding vector, pQE-60, for cloning of the Bap-5 fragments. Therefore, all constructs with the exception of pET11a(BrkA) encoded a His$_6$ tag. The cloned amplicons were subjected to restriction digestion according to Table 6. The restricted fragments were then separated by agarose gel electrophoresis and the bands of correct size purified (Figure 19b). The expression vectors, pET-11a, pET-33b or pQE-60, were digested with either *Nde*I/*BamH*I, *Bsp*HI/*BamH*I or *Nco*I/*Bgl*II as appropriate (Table 6) and gel purified. Ligations were performed according to section 2.9.2 and the products were transformed into either BL21(DE3) or M15p(Rep4) *E. coli* strains as appropriate (Table 3). The resultant colonies were checked for presence of an insert of correct size with the enzymes used to generate the cohesive termini prior to ligation. Those containing an insert of correct size were subjected to expression analysis.

Table 6: Primers and restriction sites used to generate expression constructs.

Constructs	Insert and primers used for amplification†	Enzyme sites in primers used for cloning into expression vector*
pCRII(BCT), pET11a(BCT)	BrkA C-terminus DNA, *brkACT* BRKAC-TERM5' and BRKA3'	*Nde*I, *Bam*HI
pCRII(PCT), pET33b(PCT)	Prn C-terminus DNA, *prnCT* PRNC-TERM5' and PRN3'	*Bsp*HI, *Bam*HI
pCRII(TCT), pET33b(TCT)	Tcf C-terminus DNA, *tcfCT* TCFC-TERM5' and TCF3'	*Nco*I, *Bam*HI
pCRII(NTS), pQE-60 (NTS)	Specific Bap-5 portion, *nts* NTSFOR and NTSREV	*Nco*I, *Bgl*II
pCRII(Bap5M1), pQE-60(BapM1)	Bap-5 from the 4 potential start codons NTM1 and BAP-5CTERM3'	*Bsp*HI, *Bgl*II
pCRII(Bap5M2), pQE-60(BapM2)	NTM2 and BAP-5CTERM3'	
pCRII(Bap5M3), pQE-60(BapM3)	NTM3 and BAP-5CTERM3'	
pCRII(Bap5M4), pQE-60(BapM4)	NTM4 and BAP-5CTERM3'	
pCRII(BP5CT), pQE-60(BP5CT)	Bap-5 C-terminus DNA, *Bap-5CT* BAP-5CTERM5' and BAP-5CTERM3'	*Nco*I, *Bgl*II

†See Table 5 for primer sequences. **Bsp*HI generates cohesive overhangs that are compatible with *Nco*I. This site was used in place of *Nco*I in primers that amplified a sequence containing *Nco*I.

3.2.2 Expression of recombinant autotransporter domains

The amount of protein expressed as a proportion of whole-cell protein (the expression level) was found to be comparable regardless of IPTG concentration in the range 0.2-1.5 mM (see Figure 21). The expression levels in both 2xYT and LB media were comparable for all systems tested, although 2xYT enabled greater *E. coli* cell density and therefore more recombinant protein per unit volume of culture (data not shown). A final concentration of 0.2 mM IPTG was therefore used to induce expression of all proteins. In all cases, protein expression was tightly regulated and was not detectable from uninduced cultures in either Coomassie-blue stained gels or, where performed, immunoblots with anti-His_6 antibodies (Figure 28). The expression of Bap-5 from clones pQE-60(BapM1) and pQE-60(BapM2) can be seen in Figure 22. The predicted molecular weight of Bap-5 (from the 1st ATG) is 79.5 KDa, and bands can be seen at 80, 50 and 28 KDa. These bands may represent an unprocessed form of Bap-5 (80 KDa) and a processed form consisting of a C-terminus (30 KDa) and a mature N-terminal band (50 KDa).

The pertactin C-terminus protein expressed from 3 clones containing pET33b(PCT) at different IPTG concentrations was noted previously (Figure 21). The C-terminal portions of BrkA, Tcf and Bap-5 were expressed at approximately the same level as the Prn C-terminus and the apparent molecular weights were approximately 30 KDa in all cases, as predicted. The clone pET33b(NTS) expressing the specific N-terminal portion of Bap-5 had a predicted Mw of 16.5 KDa and reproducibly produced proteins of size 16.5 KDa and 14 KDa. The 14KDa protein is probably due to proteolytic cleavage of the 16.5KDa protein (Figure 23). An insert-free control was also induced under the same conditions as the recombinant strains and showed no apparent change in whole-cell protein profile following induction (lanes 10-12, Figure 21).

Figure 21: Whole-cell lysates of BL21(DE3) *E. coli* expressing the C-terminus of pertactin under different ITPG conditions (pET33b used). Cultures were grown in 1 L of 2xYT and were induced at OD_{600} of 0.4. Note that the amount of protein loaded onto the gel was not normalised.

KDa: 10 KDa ladder

Lane 1: pET33b(PCT) immediately prior to induction (Time 0 h)

Lane 2: pET33b(PCT) without induction of recombinant expression (Time 5 h)

Lane 3: pET33b(PCT) with 0.2 mM IPTG induction (Time 5 h)

Lane 4: pET33b(PCT) immediately prior to induction (Time 0 h)

Lane 5: pET33b(PCT) without induction of recombinant expression (Time 5 h)

Lane 6: pET33b(PCT) with 0.75 mM IPTG induction (Time 5 h)

Lane 7: pET33b(PCT) immediately prior to induction (Time 0 h)

Lane 8: pET33b(PCT) without induction of recombinant expression (Time 5 h)

Lane 9: pET33b(PCT) with 1.5 mM IPTG induction (Time 5 h)

Lane 10: pET33b(No Insert) immediately prior to induction (Time 0 h)

Lane 11: pET33b(No Insert) without induction of recombinant expression (5 h)

Lane 12: pET33b(No Insert) with 1.5 mM IPTG induction (5 h)

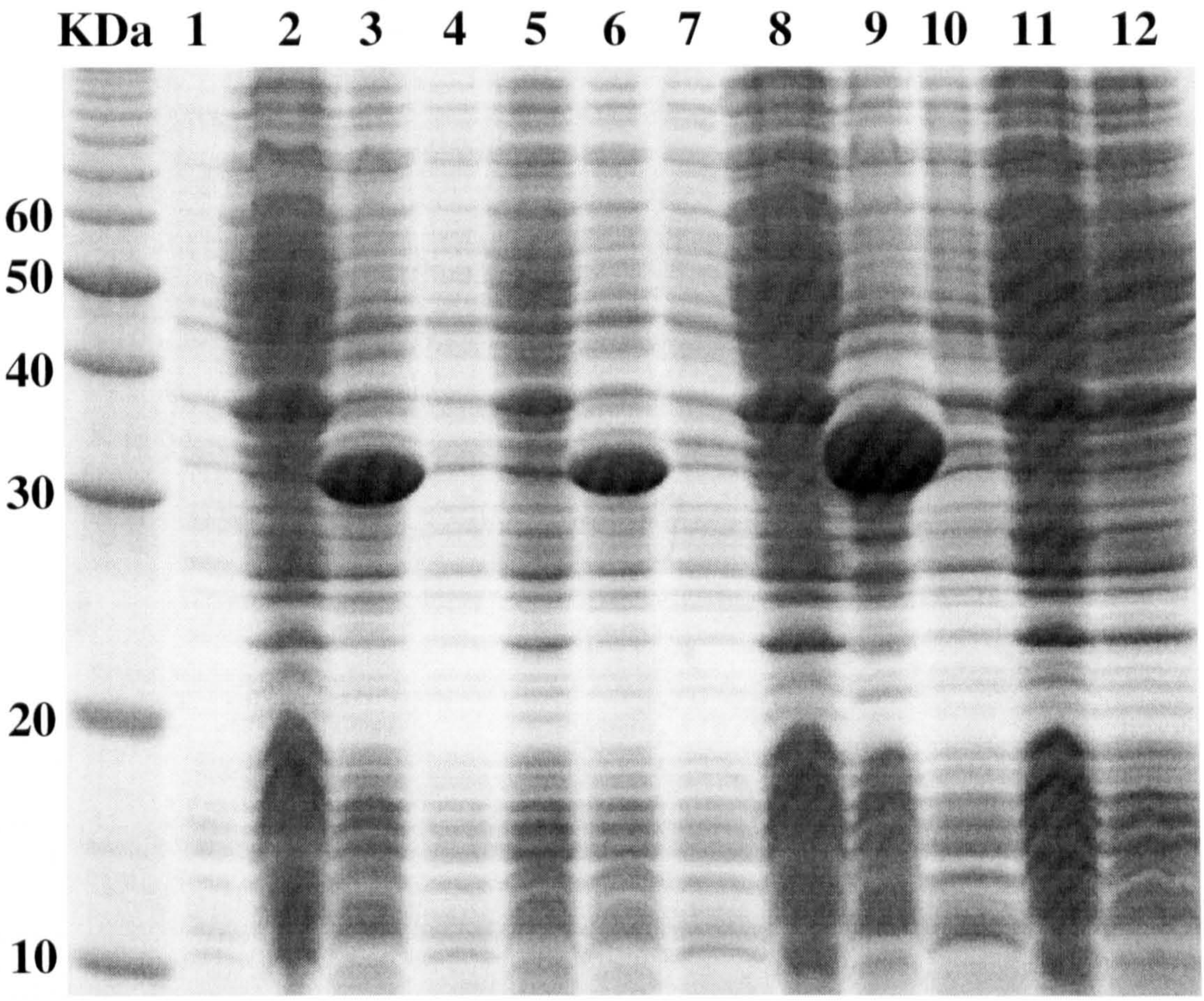

KDa
1
2
3
4
5
6
7
8
9
10
11
12
60
50
40
30
20
10

Figure 22: Whole-cell lysates of *E. coli* strain M15p(Rep4) containing Bap-5 expression constructs (pQE-60 used). Note the bands of apparent molecular weight 80 KDa, 50 KDa and 28 KDa (marked with arrows) in lanes 2 and 4.

KDa: 10 KDa ladder

Lane 1: pQE-60(Bap-5M1) without induction of recombinant protein (Time 5 h)

Lane 2: pQE-60(Bap-5M1) with 0.2 mM IPTG induction (Time 5 h)

Lane 3: pQE-60(Bap-5M2) without induction of recombinant protein (Time 5 h)

Lane 4: pQE-60(Bap-5M2) with 0.2 mM IPTG induction (Time 5 h)

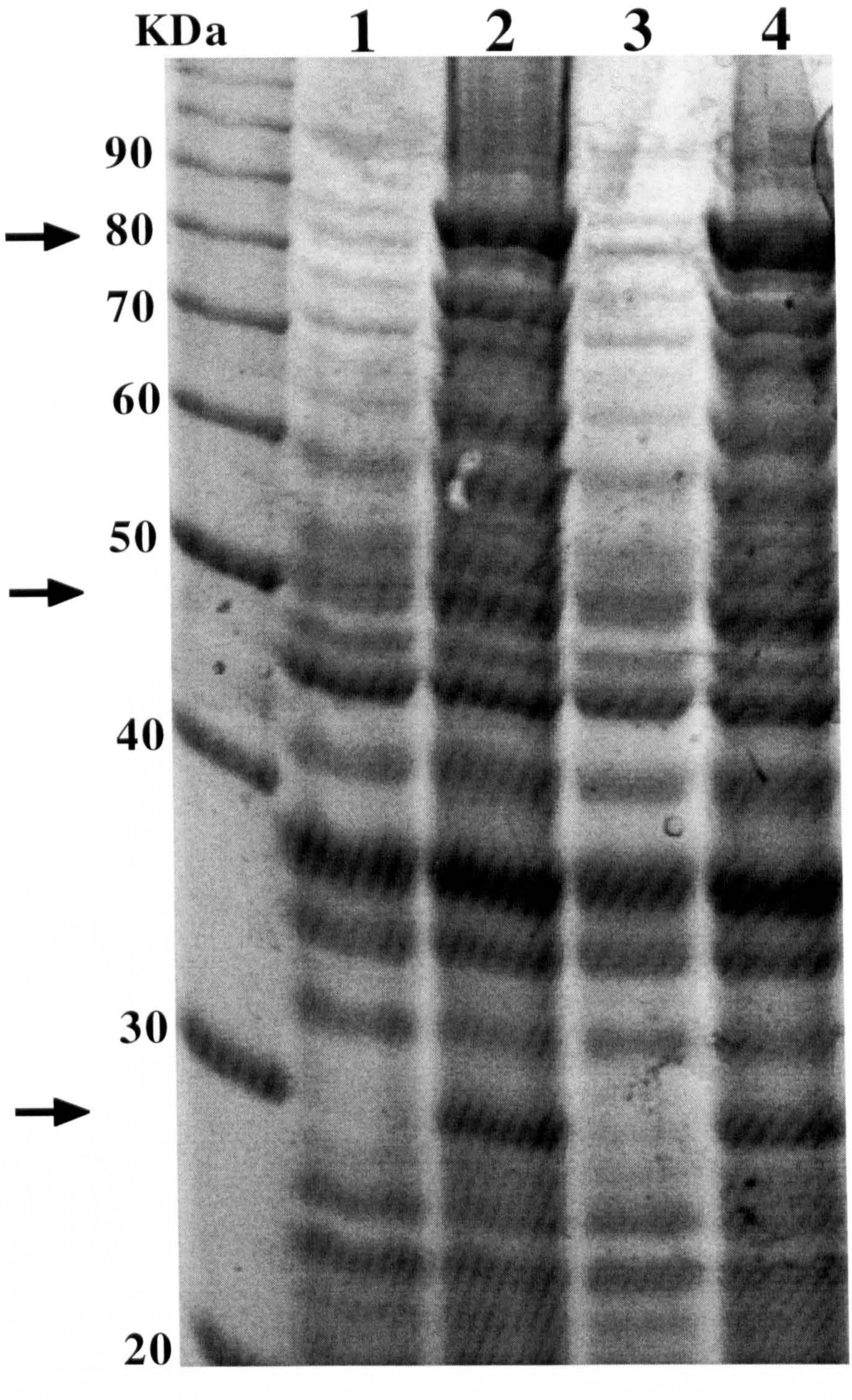
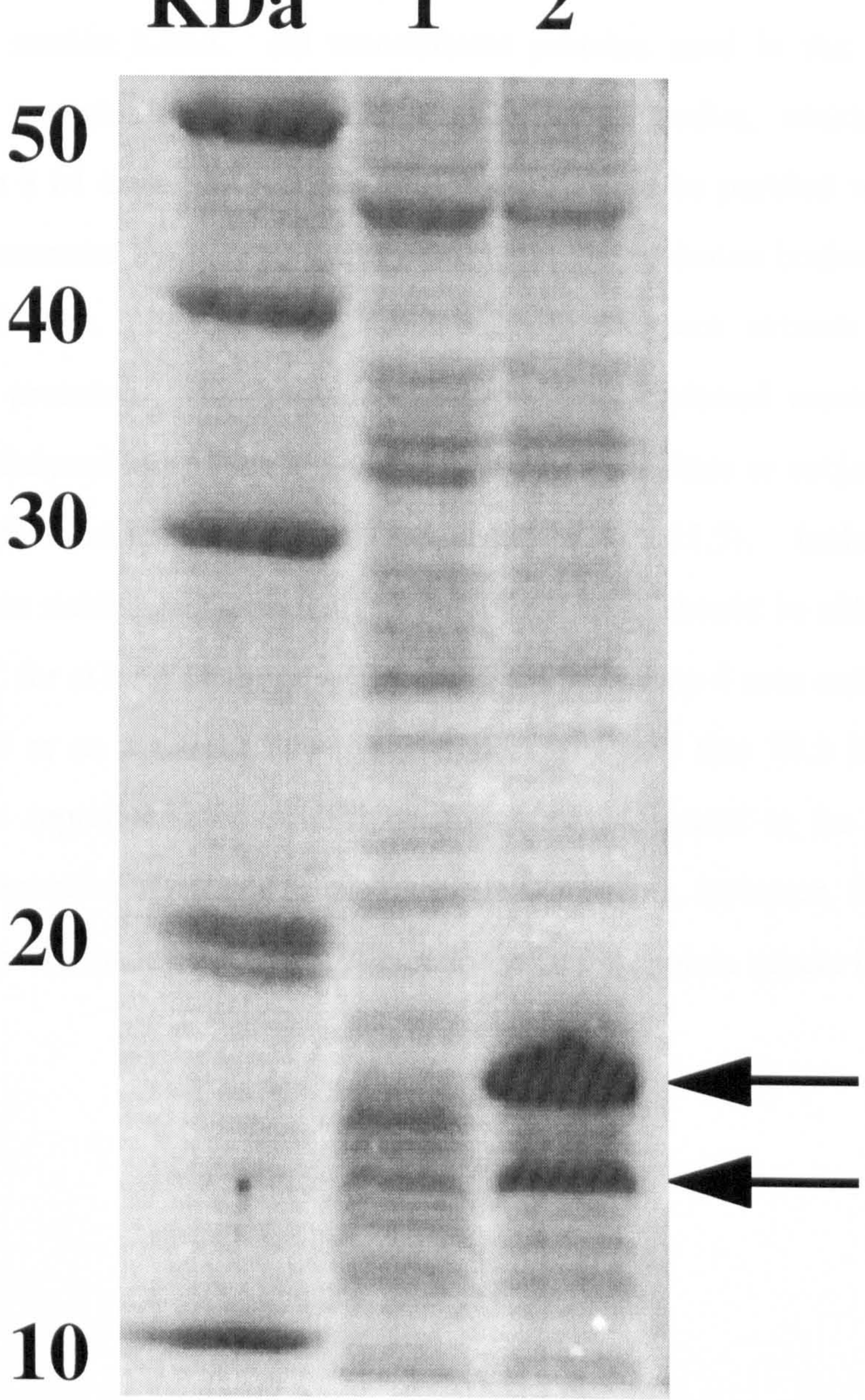

KDa
1
2
3
4
90
80
70
60
50
40
30
20

Figure 23: Whole-cell lysates of *E. coli* strain M15p(Rep4) containing an NTS expression contruct (pQE-60 used). Note the presence of bands with apparent molecular weights of 16.5 KDa and 14.5 KDa in lane 2 (see arrows).

KDa: 10 KDa ladder

Lane 1: pQE-60(NTS) without induction of recombinant protein (Time 5 h)

Lane 2: pQE-60(NTS) with 0.2 mM IPTG induction (Time 5 h)


KDa
1
2
50
40
30
20
10

3.2.3 Purification of recombinant autotransporter domains

3.2.3.1 Partial purification by urea extraction

To partially purify the recombinant proteins, the urea soluble fraction was prepared according to section 2.15.3. All recombinant proteins used in this study could be extracted as insoluble protein, probably as inclusion bodies, which could then be solubilised in 8 M urea. All recombinant proteins could be purified to approximately 95% purity (assessed by SDS-PAGE) by extracting the inclusion bodies and solubilising them in 8M urea. Figures 24a and 24b show the urea extracts of the various recombinant proteins. This provided semi-purified denatured recombinant proteins which were dialysed into either 1M urea in PBS or PBS alone or subjected to Ni-NTA affinity purification as appropriate (sections 2.15.4, 2.15.5). Inclusion bodies are generally more stable against proteolysis and all proteins should be able to be stored in urea at -20°C for at least 18 months (Enfors 1992). The Bap-5 urea extract (Figure 24b) shows Bap-5 at an apparent Mw of 90 KDa (predicted size 79.5 KDa). This size anomaly was reproduced on several occasions and appeared to be gel specific and occurred irrespective of sample preparation. It is possible, however, that the excessive amount of protein loaded in Figure 24b made the Bap-5 protein appear larger.

Figure 24a: Whole-cell lysates and urea-soluble fraction of recombinant *E. coli* induced with 0.2 mM IPTG to express Prn (from pET33b), BrkA (from pET33b), Tcf (from pET33b) and Bap-5 C-terminal domains (from pQE-60) and the Bap-5 specific domain (NTS)) (from pQE-60). The major bands represent the recombinant domains.

Lane 1: Whole-cell lysate of *E. coli* expressing PCT

Lane 2: Urea soluble fraction of *E.coli* expressing PCT

Lane 3: Whole-cell lysate of *E. coli* expressing TCT

Lane 4: Urea soluble fraction of *E.coli* expressing TCT

Lane 5: Whole-cell lysate of *E. coli* expressing BCT

Lane 6: Urea soluble fraction of *E.coli* expressing BCT

Lane 7: Whole-cell lysate of *E. coli* expressing Bap-5CT

Lane 8: Urea soluble fraction of *E.coli* expressing Bap-5CT

Lane 9: Whole-cell lysate of *E. coli* expressing NTS

Lane 10: Urea soluble fraction of *E.coli* expressing NTS

Figure 24b: Urea-soluble fraction of *E. coli* strain M15p(Rep4) containing pQE-60(Bap-5M1) and expressing Bap-5 from the 1st potential start codon can be seen in lane 1.

Figure 24a

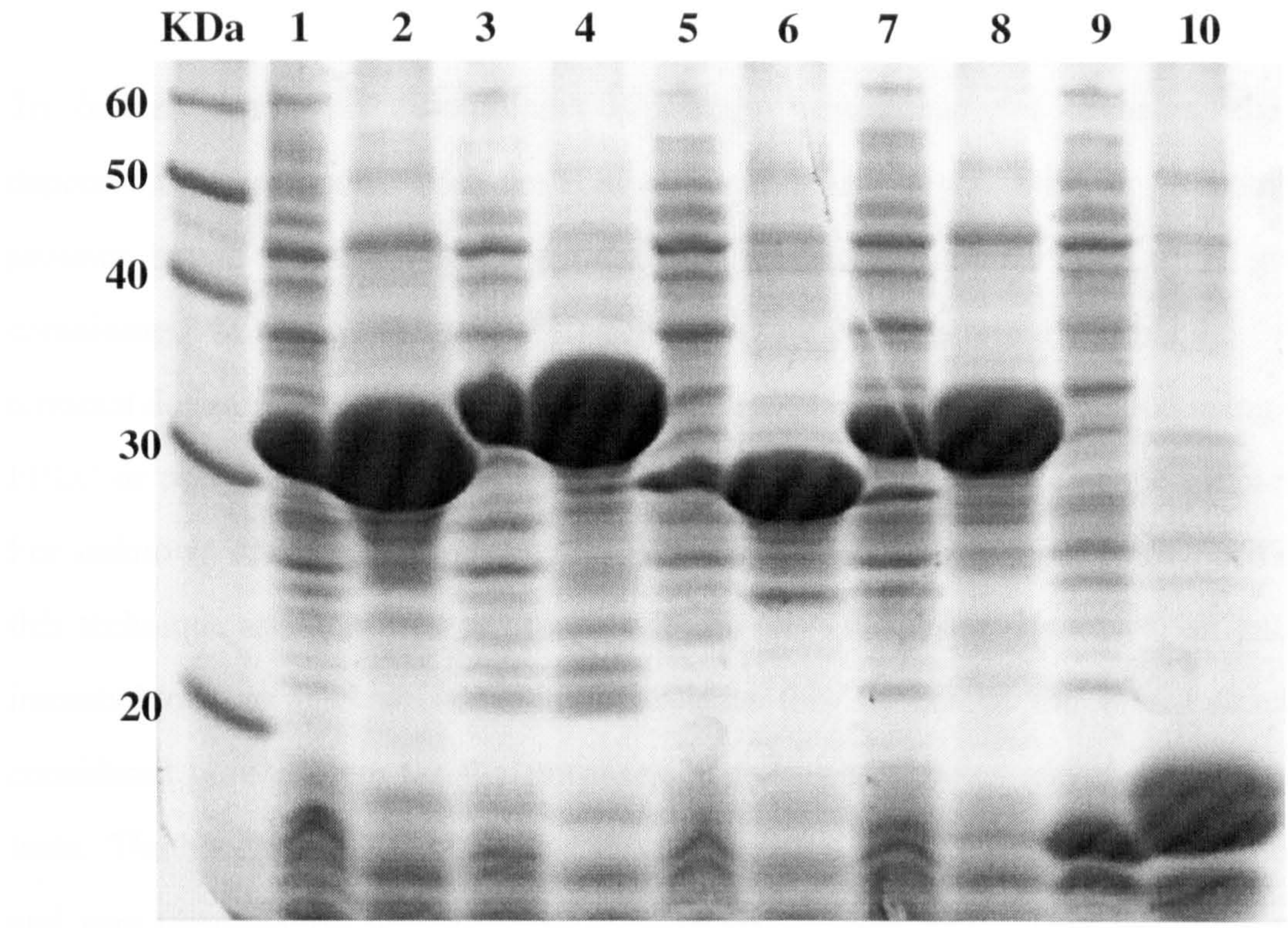
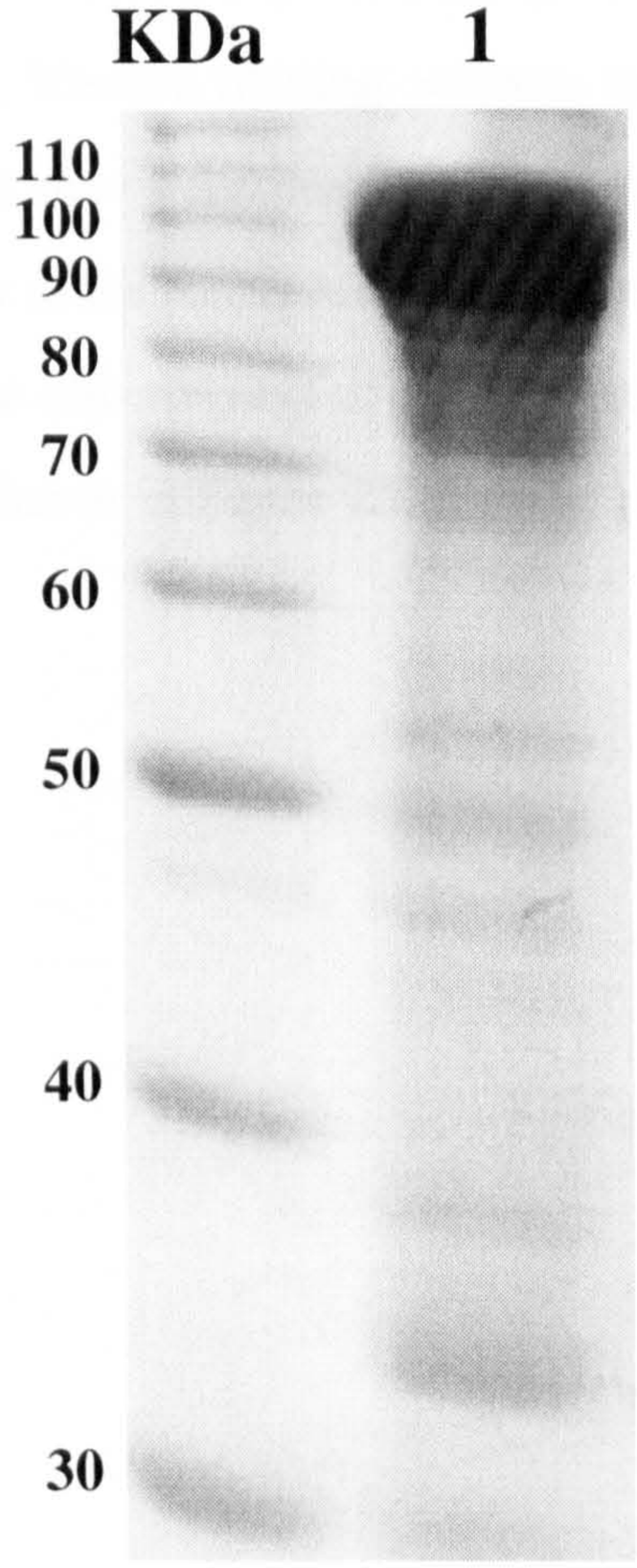


Figure24b



3.2.3.2 Purification by affinity chromatography

To further purify the recombinant His_6-tagged autotransporter domains, the pH-dependent interaction between the His_6 tags and Ni-NTA was utilised. Recombinant proteins bound to the column and were then washed and eluted using a pH gradient containing 8 M urea. The optimal elution pH for all proteins tested was pH 4.5. All C-terminal domains and the NTS domain were purified by this affinity method using either FPLC or using Ni-NTA microspin columns depending on the scale of the preparation. For unknown reasons, in a preliminary experiment, whole Bap-5 was not purified by this technique and appeared not to bind to the Ni-NTA resin. However, an anti-His_6 immunoblot was not performed to investigate this further as the urea extract was considered pure enough for the purposes of raising antibodies and mouse protection tests. The band of 14 KDa obtained from expression of NTS did not bind to the resin and was presumably, therefore, not His_6-tagged (Figure 27). This assumption was confirmed, as shown in Figure 28, as the lower band (14.5 KDa) did not bind anti-His_6 monoclonal antibodies during Western blotting, whereas the larger 16.5 KDa band did bind the antibody. Figure 25 shows the purification of the Prn C-terminus. On this occasion, the binding capacity of the Ni-NTA has been exceeded (see major protein band in lanes 2 and 3), although the effectiveness of the purification was not reduced. Only representative fractions are shown in Figure 25. Similar profiles were obtained for the other C-terminal proteins.

Figure 25: FPLC Purification of the pertactin C-terminus using Ni-NTA affinity. The binding capacity of the Ni-NTA was exceeded, as can be seen by the presence of protein of apparent Mw 32 KDa in the flow-through. Also, more protein could potentially have been eluted in Buffer E, as the recombinant protein is still bound to the column and is present the Buffer F eluate.

KDa: 10 KDa ladder

Lane 1: Urea extract of pertactin C-terminus loaded onto column

Lane 2: Buffer B, first flow through

(8 M urea, 0.1 M sodium phosphate, 0.01 M Tris, pH 8.0)

Lane 3: Buffer C, eluate 1

(8 M urea, 0.1 M sodium phosphate, 0.01 M Tris HCl, pH 6.3)

Lane 4: Buffer C, eluate 2

Lane 5: Buffer D, eluate 1

(8 M urea, 0.1 M sodium phosphate, 0.01 M Tris HCl, pH 5.9)

Lane 6: Buffer E, eluate 1

(8 M urea, 0.1 M sodium phosphate, 0.01 M Tris HCl, pH 4.5)

Lane 7: Buffer F, eluate 1

(6 M guanidine hydrochloride, 0.2 M acetic acid)

Lane 8: Buffer F, eluate 2

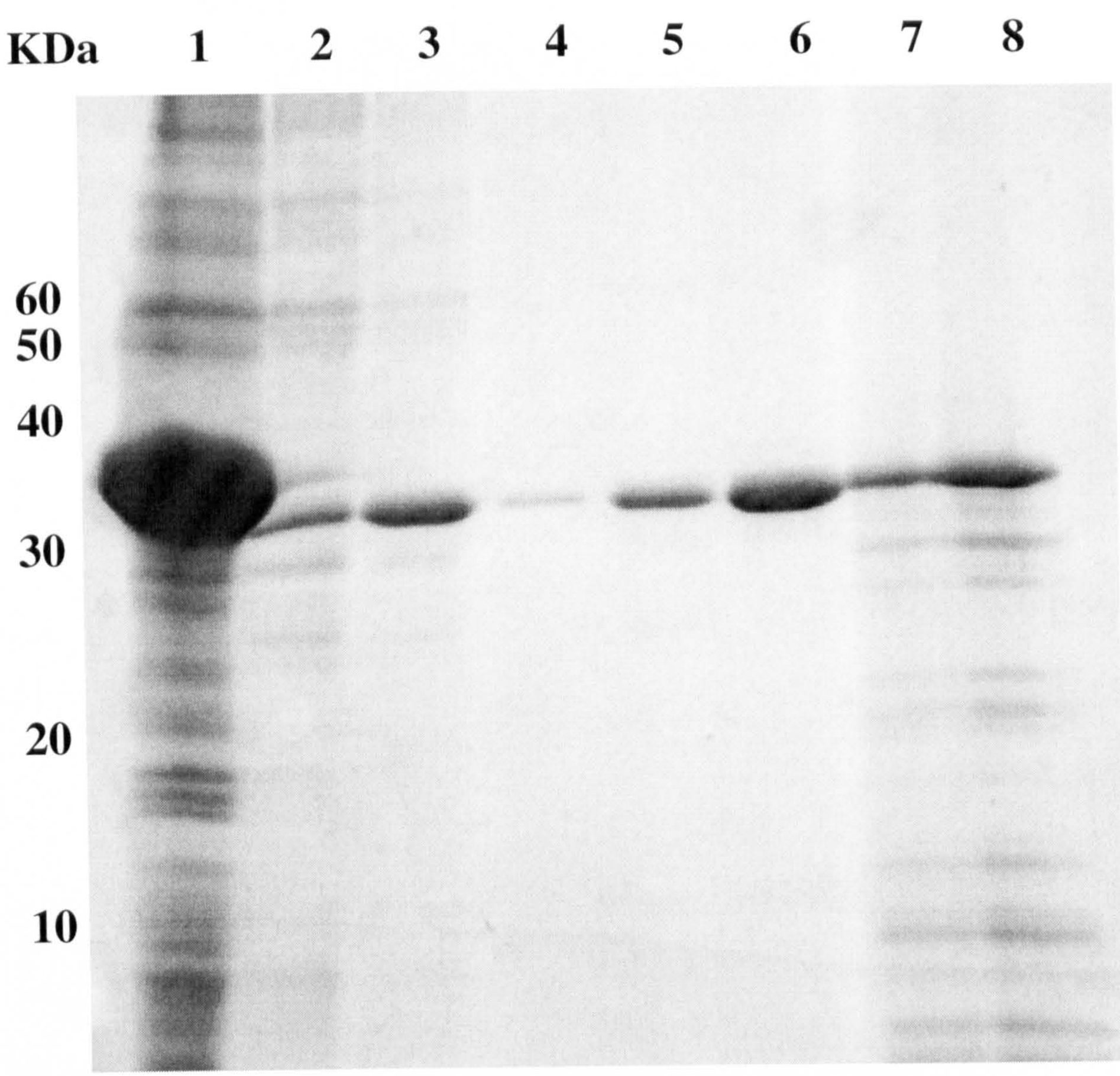

KDa
1
2
3
4
5
6
7
8
60
50
40
30
20
10

Figure 26: Purification of the Bap-5 C-terminus using Ni-NTA microspin columns.

KDa: 10 KDa ladder

Lane 1: Urea extract of pertactin C-terminus as loaded onto column

Lane 2: Buffer C, eluate 1

(8 M urea, 0.1M sodium phosphate, 0.01M Tris Hcl, pH 6.3)

Lane 3: Buffer C, eluate 2

(8 M urea, 0.1M sodium phosphate, 0.01M Tris Hcl, pH 6.3)

Lane 4: Buffer C, eluate 3

(8 M urea, 0.1M sodium phosphate, 0.01M Tris Hcl, pH 6.3)

Lane 5: Buffer E, eluate 1

(8 M urea, 0.1M sodium phosphate, 0.01M Tris, pH 4.5)

Lane 6: Buffer E, eluate 2

(8 M urea, 0.1M sodium phosphate, 0.01M Tris, pH 4.5)

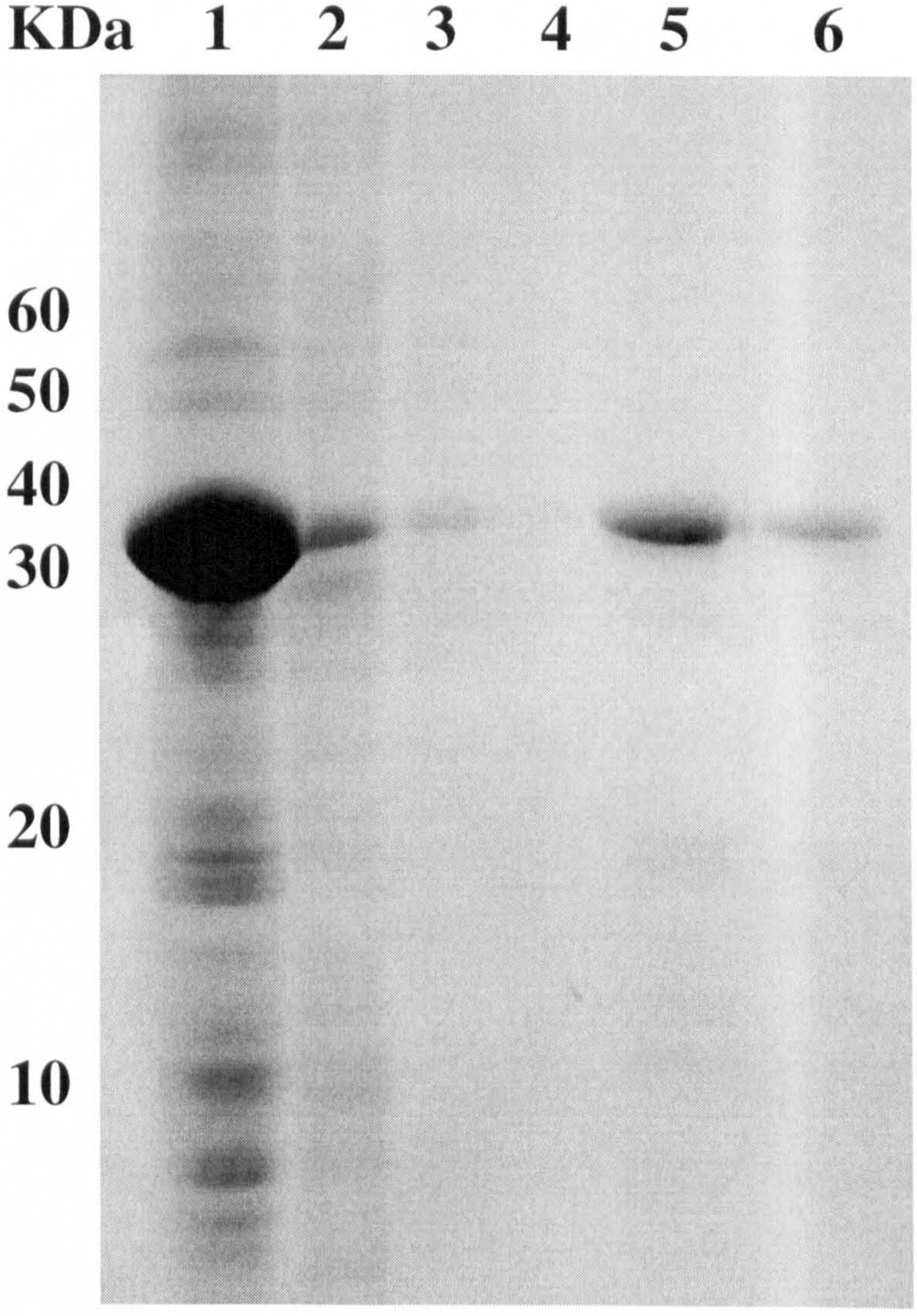

KDa
1
2
3
4
5
6
60
50
40
30
20
10

Figure 27: Purification of the Bap-5-specific NTS domain by FPLC Ni-NTA affinity chromatography (solid arrow, lane 8). The 14.5 KDa band is present in Lane 1 and does not, therefore, appear to bind to the resin (dashed arrrow). This suggests that this protein is not His_6-tagged.

KDa: 10 KDa Protein Ladder

Lane 1: Buffer B, first flow through

Lane 2: Buffer B, second flow through

Lane 3: Buffer B, third flow through

Lane 4: Buffer C, eluate 2

Lane 5: Buffer C, eluate 3

Lane 6: Buffer D, eluate 2

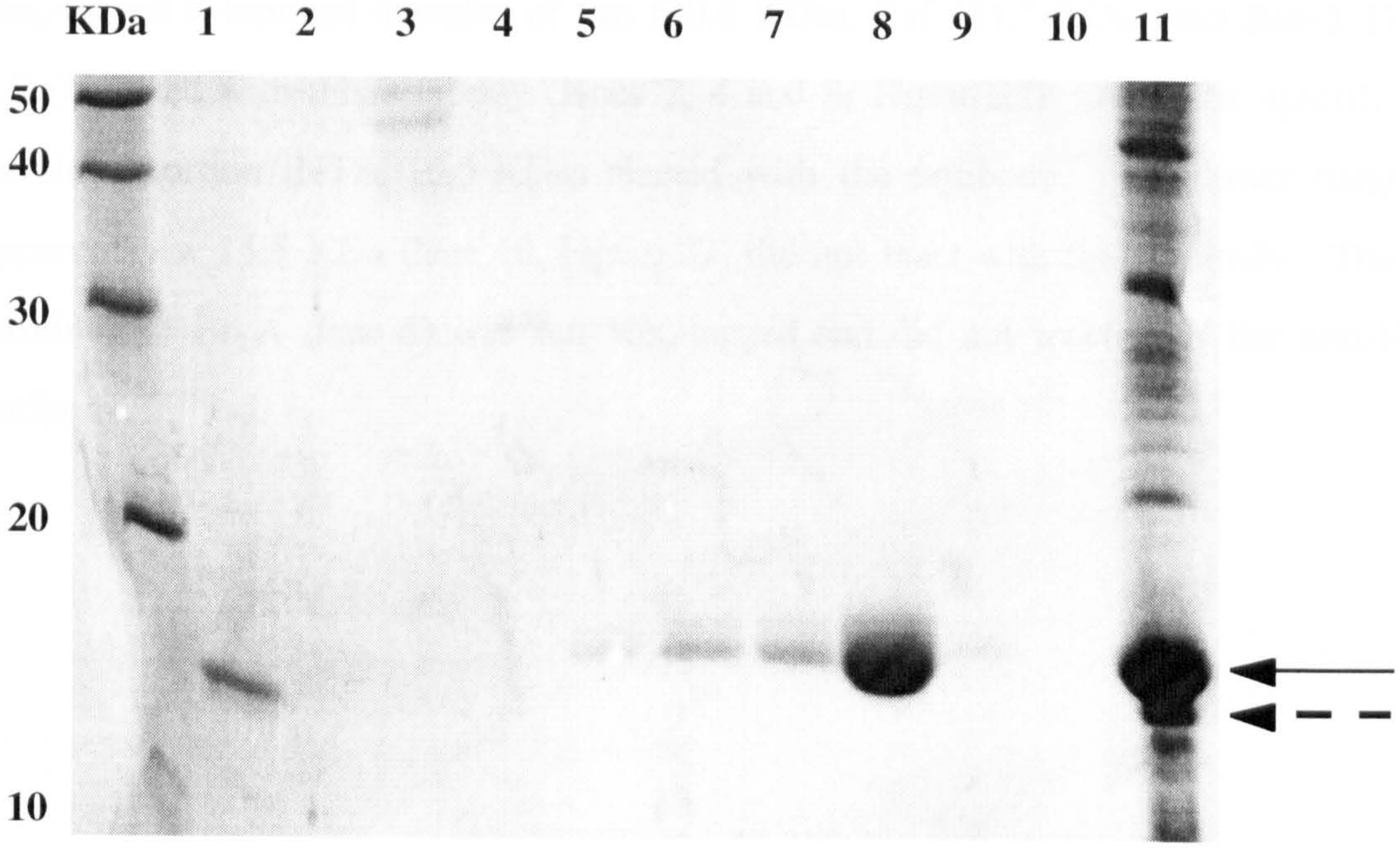
Lane 7: Buffer E, eluate 1

Lane 8: Buffer E, eluate 2

Lane 9: Buffer E, eluate 3

Lane 10: Gap-no sample loaded

Lane 11: Urea extract of NTS loaded onto column


KDa
1
2
3
4
5
6
7
8
9
10
11
50
40
30
20
10

3.2.4 Immunoblotting of recombinant domains with anti-His_6 antibody

To confirm that a His_6 tag was present on the predicted recombinant proteins and to confirm the lack of expression in non-induced clones, whole-cell lysates were immunoblotted with an anti-His_6 antibody. As predicted from the construct design, the His_6-tagged C-terminal domains of Prn (30.6 KDa), Tcf (31.7 KDa) and Bap-5 (31.3 KDa) reacted with this antibody (lanes 2, 4 and 8; Figure 27). Also the specific N-terminus portion (NTS, 16.5 KDa) reacted with the antibody. The lower band of apparent Mw 14.5 KDa (lane 10, Figure 27) did not react with this antibody. The C-terminus of BrkA (lane 6) was not His_6-tagged and did not react with the anti-His_6 antibody.

Figure 28: Immunoblot of induced and non-induced whole-cell lysates using anti-His_6 monoclonal antibodies. There appears to be some degradation of the recombinant proteins, which resulted in multiple bands.

KDa: 10 KDa protein ladder

Lane 1: Whole-cell lysate of *E. coli* not induced to express PCT

Lane 2: Whole-cell lysate of *E. coli* expressing PCT

Lane 3: Whole-cell lysate of *E. coli* not induced to express TCT

Lane 4: Whole-cell lysate of *E. coli* expressing TCT

Lane 5: Whole-cell lysate of *E. coli* not induced to express BCT

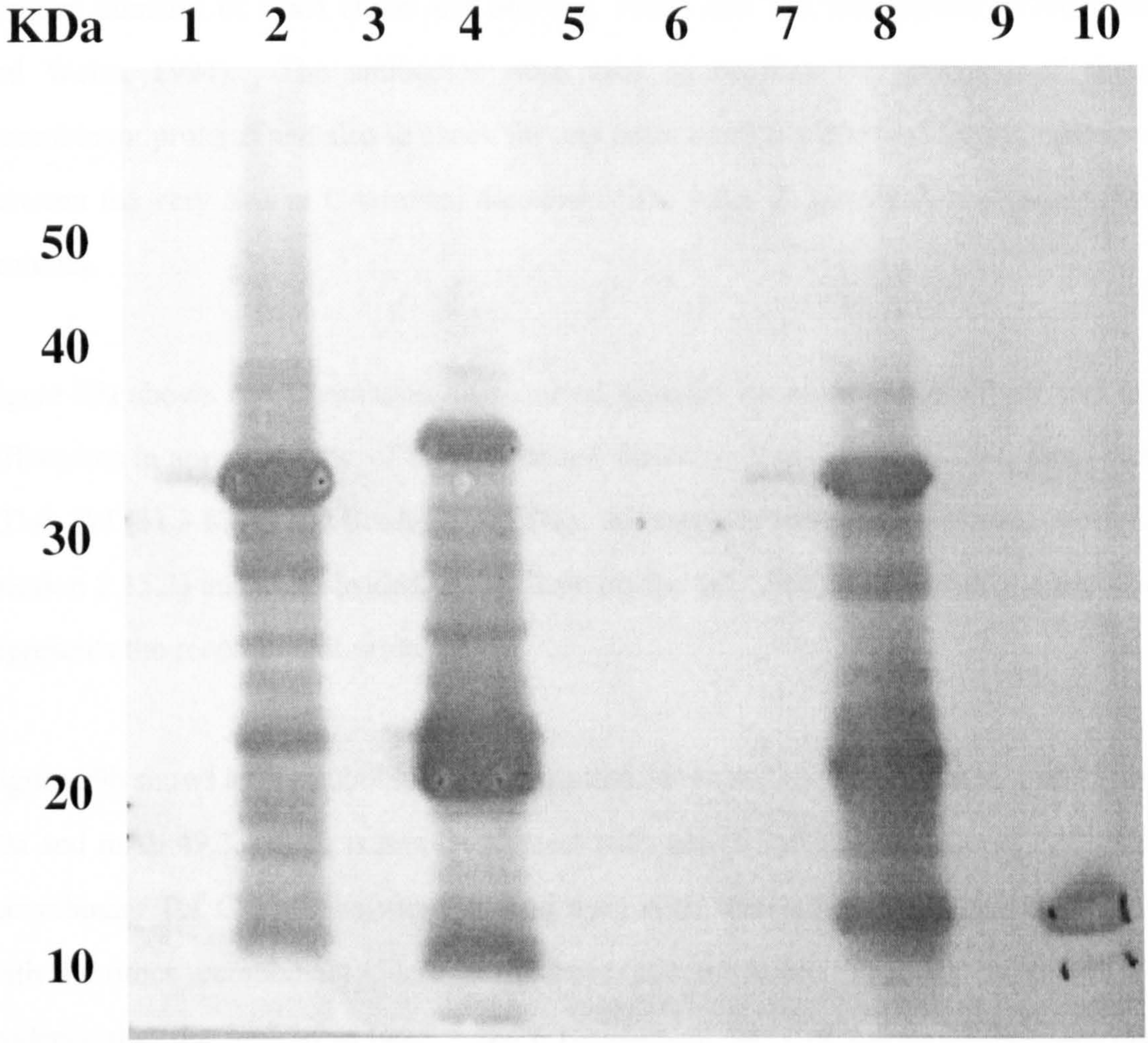
Lane 6: Whole-cell lysate of *E. coli* expressing BCT

Lane 7: Whole-cell lysate of *E. coli* not induced to express Bap-5CT

Lane 8: Whole-cell lysate of *E. coli* expressing Bap-5CT

Lane 9: Whole-cell lysate of *E. coli* not induced to express NTS

Lane 10: Whole-cell lysate of *E. coli* expressing NTS


KDa
1
2
3
4
5
6
7
8
9
10
50
40
30
20
10

3.2.5 Immunoblotting of recombinant domains with mAbs

Monoclonal antibodies had been raised previously that reacted with virulence (Bvg)-regulated 28 KDa and 30 KDa bands present in the outer membrane of *B. pertussis* (Hertz and Parton, unpublished). These bands later proved to be the processed C-terminal domains of BrkA (Finn and Stevens, 1995) and Tcf respectively (Fernandez and Weiss, 1994). The antibodies were used to confirm the presence of these recombinant proteins and also to check for any cross reactivity due to common epitopes between the very similar C-terminal domains of the other *B. pertussis* autotransporter proteins.

Figure 29a shows the Coomassie blue-stained samples on an SDS-PAGE gel and the differences in apparent Mw of the C-terminal domains Bap-5 (31.3 KDa), Prn (30.6 KDa), Tcf (31.3 KDa) and BrkA (30.4 KDa). All samples were urea extracted proteins (section 2.15.3) and were loaded in duplicate on the gel. The major band in each lane represents the recombinant protein.

Figure 29b shows an immunoblot of the reaction between the samples shown in Figure 29a and mAb 49.3, which is known to react with native and SDS-denatured Tcf. The recombinant Tcf C-terminus was found to react with this antibody but cross reaction with the other recombinant C-terminal domains did not occur. This was confirmatory evidence that the protein in lanes 5 and 6 represented the C-terminus of Tcf and that Bap-5 was not antigenically closely related to Tcf.

Figure 29c shows an immunoblot with mAb 149 and the C-terminus extracts decribed above. This antibody also reacts with native Tcf C-terminus and again reacted with the recombinant form. Under the conditions used, there also appeared to be some very slight cross reactivity with Prn C-terminus and the C-terminus of Bap-5. No reaction was evident with the C-terminus of BrkA. This provides further evidence that the major protein in lanes 5 and 6 represents the C-terminus of Tcf.

Figure 29d shows an immunoblot between the aforementioned recombinant C-terminal domains with P.28 mAb, which reacts with native BrkA. This antibody did react strongly with the recombinant BrkA and, under the conditions used, also reacted very weakly with the other C-terminal domains. This cross reaction may be non-specific as many other bands, presumably *E. coli* proteins that are present in the extract, also reacted weakly with this mAb. The results, however, indicated that the major band in lanes 7 and 8 corresponds to the the BrkA C-terminus and that Bap-5 was not immunologically related to BrkA.

Figure 29a: SDS-PAGE of inclusion bodies containing recombinant C-terminal domains solubilised in urea.

KDa: 10 KDa protein ladder

Lanes 1 and 2: Bap-5 C-terminus

Lanes 3 and 4: Prn C-terminus

Lanes 5 and 6: Tcf C-terminus

Lanes 7 and 8: BrkA C-terminus

Figure 29b: Immunoblot of fractions shown in figure 29a with monclonal antibody 49.3 which reacts with native Tcf C-terminus.

KDa: 10 KDa protein ladder

Lanes 1 and 2: Bap-5 C-terminus

Lanes 3 and 4: Pertactin C-terminus

Lanes 5 and 6: Tcf C-terminus

Lanes 7 and 8: BrkA C-terminus

Figure 29a

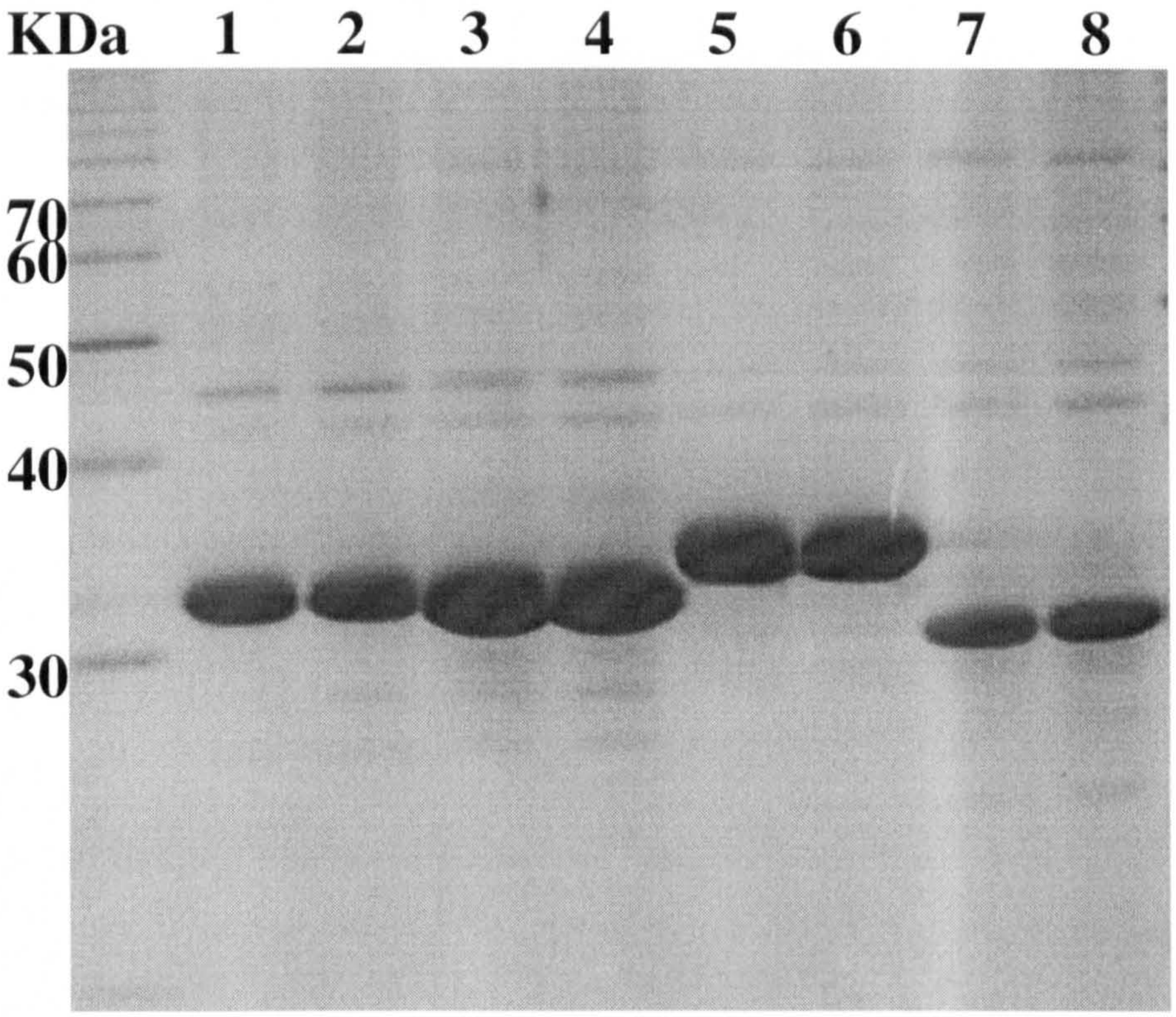


Figure 29b

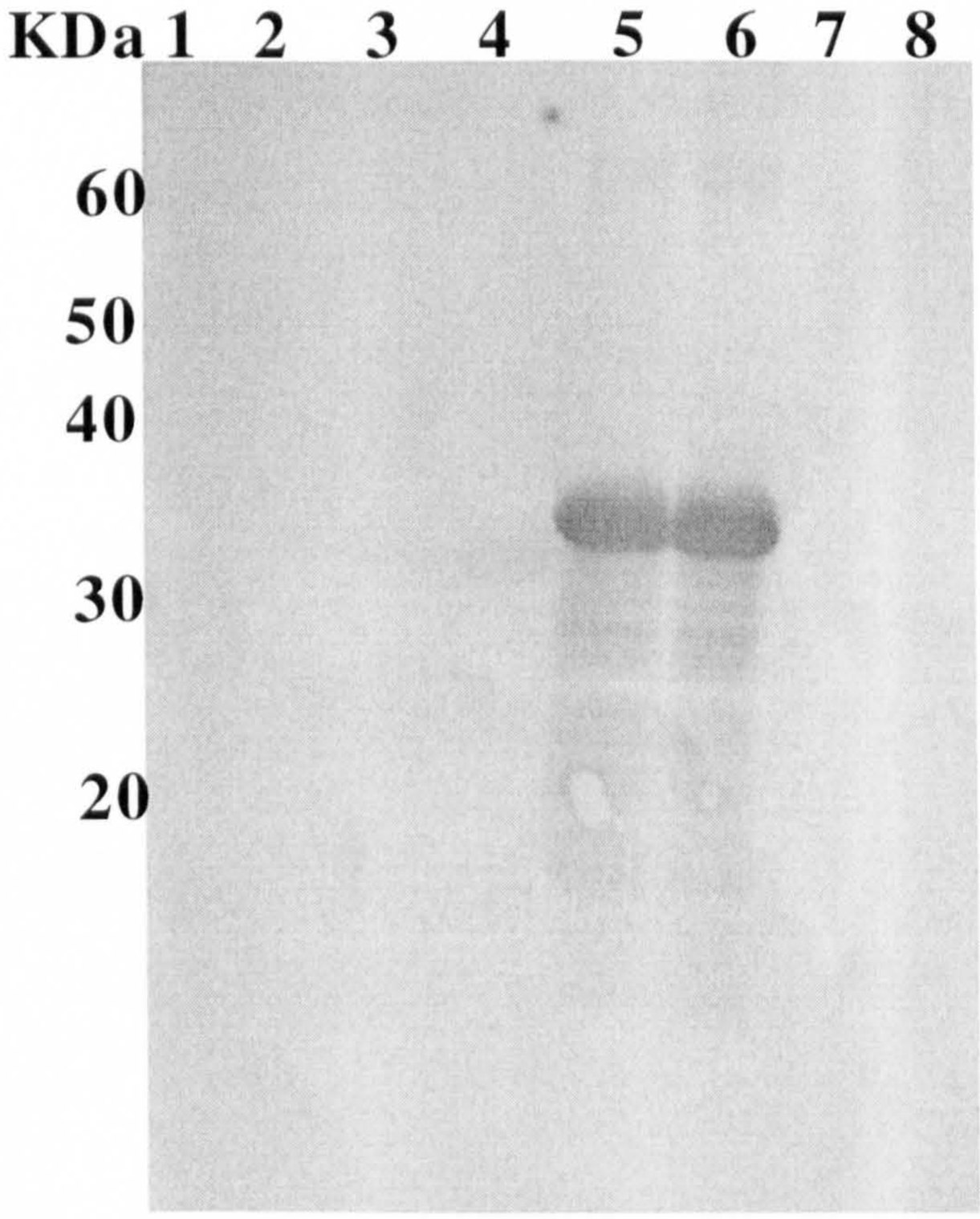


Figure29c: Immunoblot of fractions shown in figure 29a with monclonal antibody 149 which reacts with native Tcf C-terminus

KDa: 10 KDa protein ladder

Lanes 1 and 2: Bap-5 C-terminus

Lanes 3 and 4: Prn C-terminus

Lanes 5 and 6: Tcf C-terminus

Lanes 7 and 8: BrkA C-terminus

Figure 29d: Immunoblot of fractions shown in figure 29a with monclonal antibody P. 28 which reacts with native BrkA C-terminus.

KDa: 10 KDa protein ladder

Lanes 1 and 2: Bap-5 C-terminus

Lanes 3 and 4: Pertactin C-terminus

Lanes 5 and 6: Tcf C-terminus

Lanes 7 and 8: BrkA C-terminus

Figure 29c

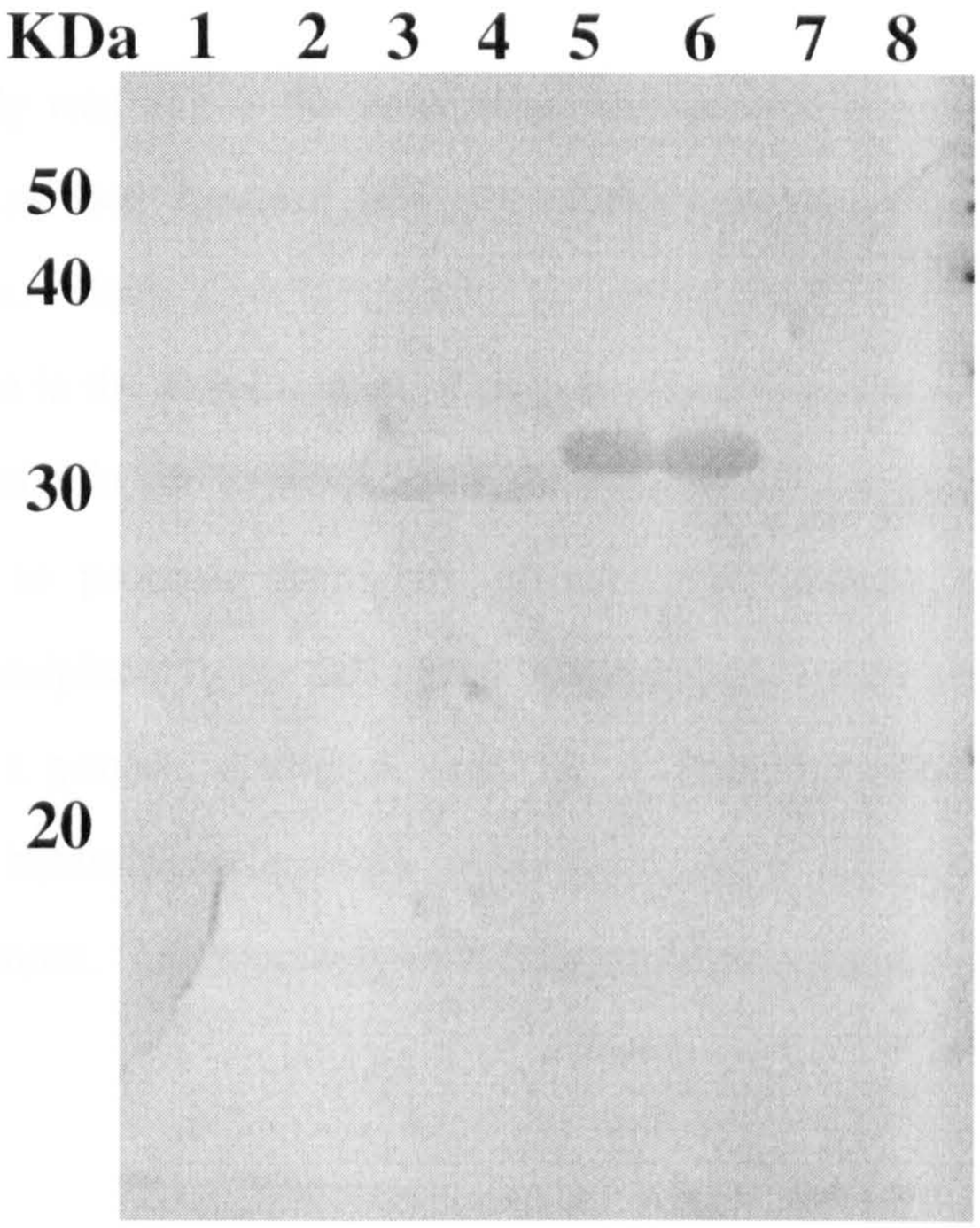
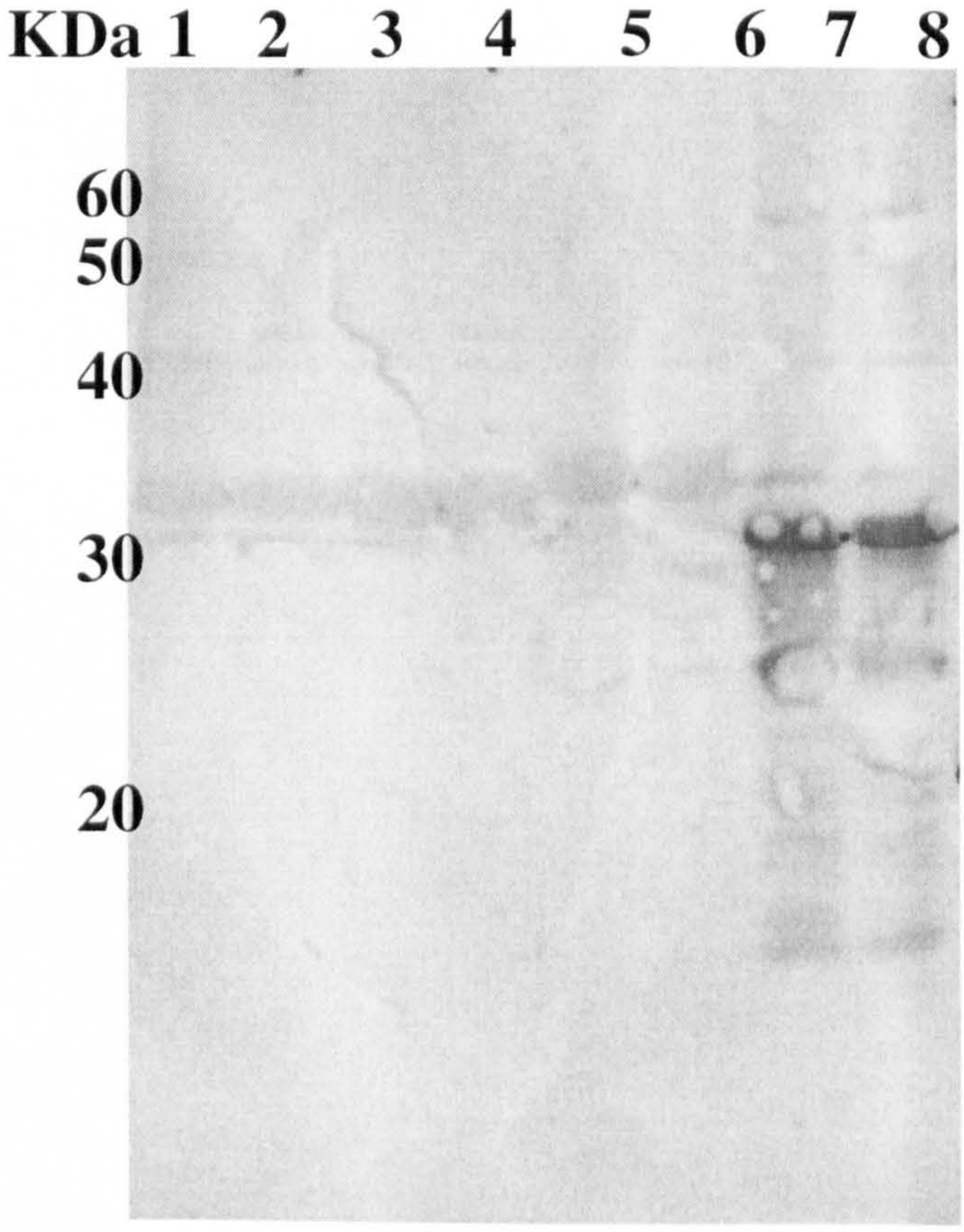


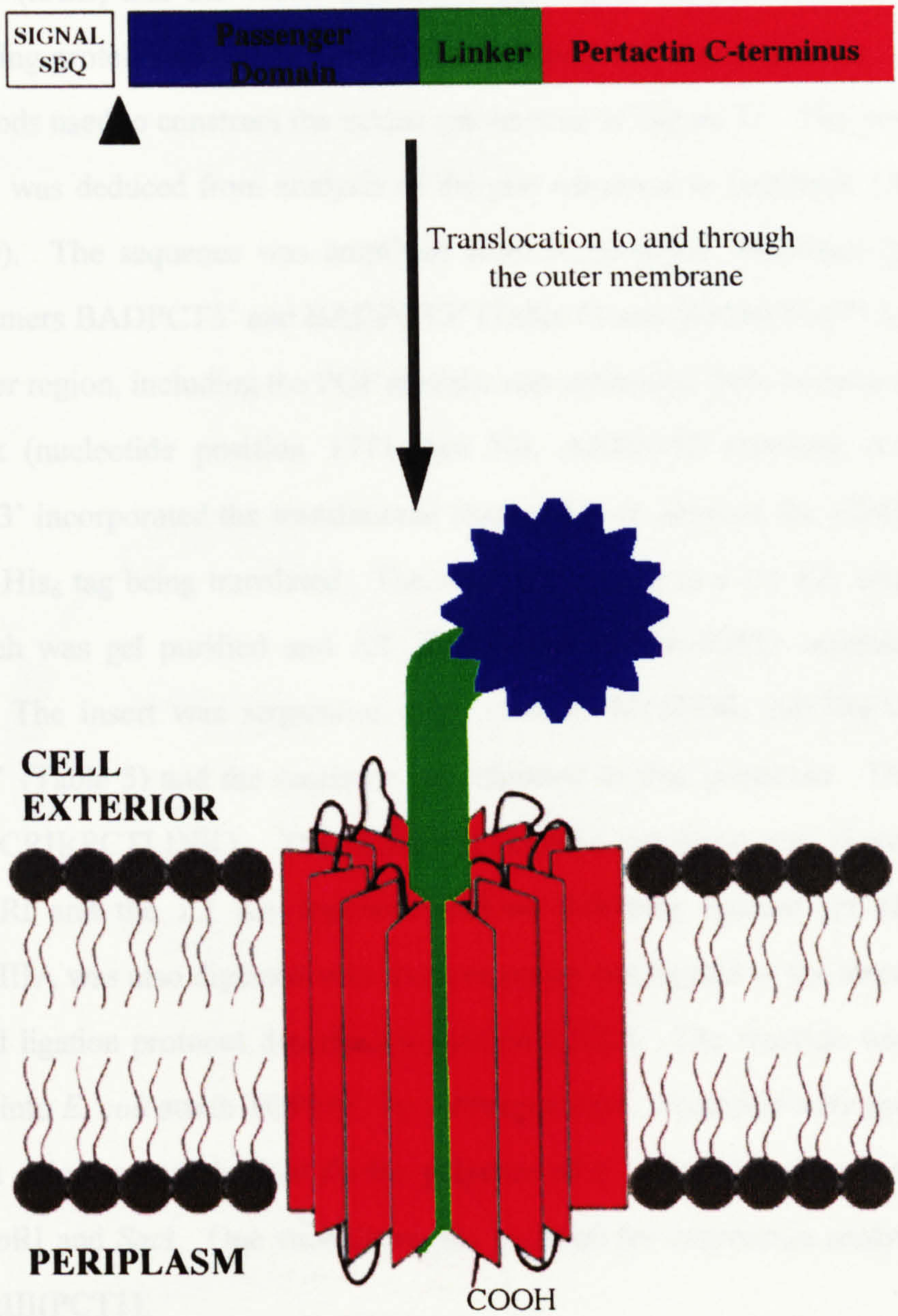
Figure 29d



3.3 Recombinant expression of BADPCT

The aim of this study was to use the autotransporter protein system to create a vector that would enable surface location and potentially surface release of heterologous proteins or peptides. Such a vector would be self-replicative and self-contained and may have great value in the development of antigen display systems for use in vaccines and may also be useful in the expression of proteins which are toxic for the host cell. The intention was to promote transport of such recombinant proteins (passenger domains) from the periplasm to the cell exterior using the C-terminal domain of pertactin and to incorporate a portion upstream from the C-terminus (linker region) to allow presentation of the heterologous protein to the pore and to allow translocation across the outer cell membrane. The topology of this heterologous display system is shown in Figure 30.

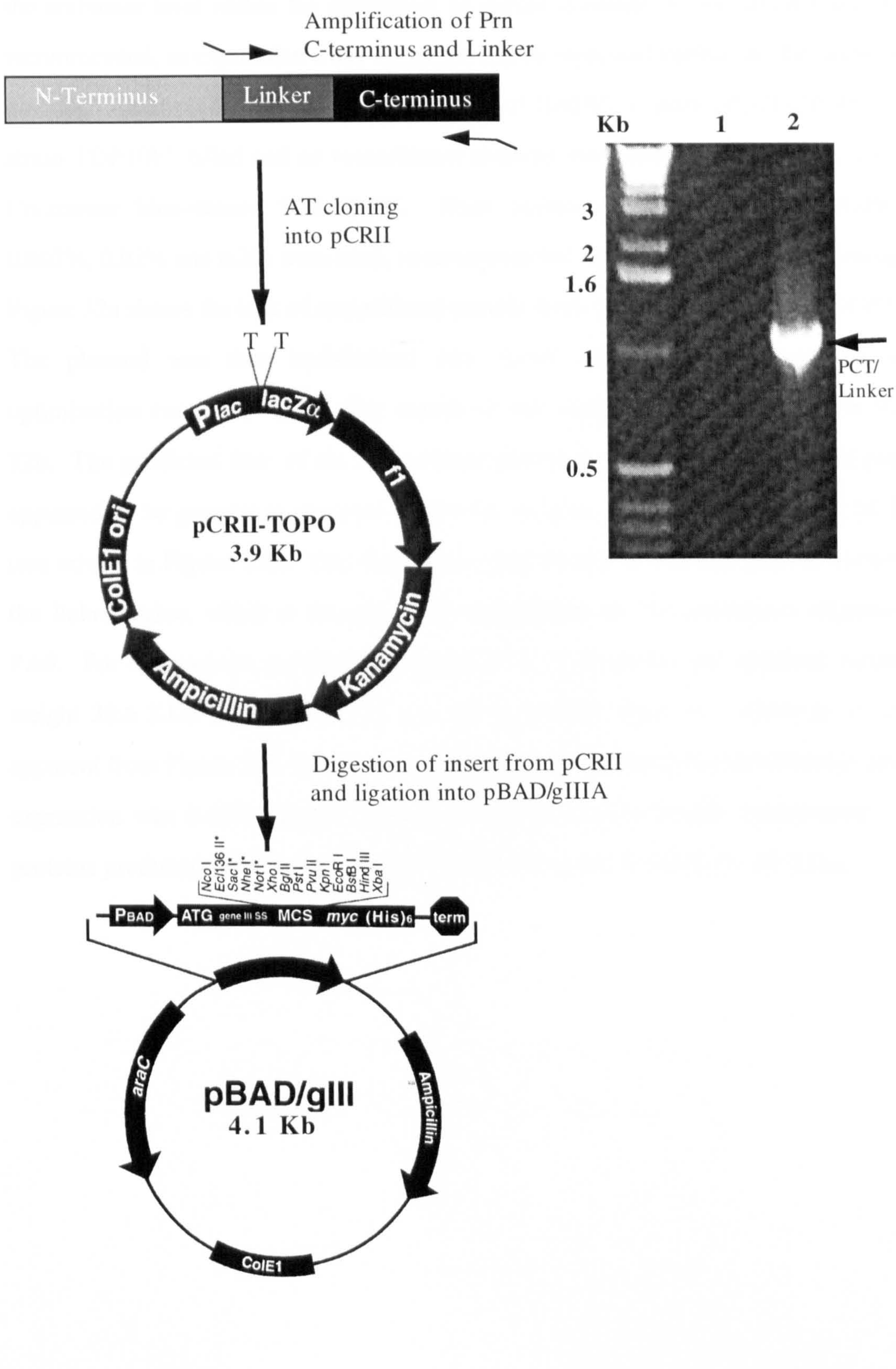
Figure 30: Topology of BADPCT heterologous antigen display. The signal sequence is removed during protein export.


SIGNAL SEQ
Passenger Domain
Linker
Pertactin C-terminus
Translocation to and through the outer membrane
CELL EXTERIOR
PERIPLASM
COOH

3.3.1 Construction of a heterologous antigen display vector

The pBAD/gIII vector was chosen as expression is tightly regulated by an arabinose promoter (*araB*) and the vector encodes a gIII signal sequence which is capable of transporting proteins via the Sec-dependent pathway into the periplasm. An outline of the methods used to construct the vector can be seen in Figure 31. The *prn/linker* target sequence was deduced from analysis of the *prn* sequence in GenBank (Accession No. AJ04560). The sequence was amplified from *B. pertussis* Taberman genomic DNA using primers BADPCT5' and BADPCT3' (Table 5) and HotstarTaq™ (section 2.7.2). The linker region, including the PQP repeats, was estimated from sequence deposited at GenBank (nucleotide position 1771 Acc No. AJ006152 encoding AANGNGQ...). BadPCT3' incorporated the translational stop codon to prevent the pBAD/gIII vector-encoded His_6 tag being translated. The reaction resulted in a 1.1 Kb amplicon (Figure 31) which was gel purified and AT cloned into pCRII-TOPO according to section 2.9.1.1. The insert was sequenced using primers M13FOR, M13REV and PRNC-TERM5' (Table 5) and the sequence was identical to that predicted. This vector was named pCRII(PCTLINK). The pCRII(PCTLINK) construct was digested with *Sac*I and *Eco*RI and the 1.1 Kb fragment purified following agarose gel electrophoresis. pBAD/gIIIA was also digested with these enzymes and ligated to the insert according to the rapid ligation protocol described in section 2.9.2.3. The reaction was transformed initially into *E. coli* strain TOP10F' by electroporation. Plasmids were purified from 20 resultant colonies and checked for the presence of a 1.1 Kb insert by double digestion with *Eco*RI and *Sac*I. One such clone was selected for expression analysis and named pBAD/gIII(PCT1).

Figure 31: A schematic overview of the construction of pBAD/gIII(PCT1). Lane 1 is a template-free negative control; Lane 2 shows the amplicon obtained.



3.3.2 Expression of BADPCT

The *E. coli* expression strains TOP10F' and LMG194 provided by the manufacturer are capable of transporting arabinose into the cell yet not metabolising it. This allows the arabinose level within the cytoplasm to remain constant. *E. coli* strain LMG194 is recommended, as expression from this strain can be repressed further by the addition of glucose. Attempts to optimise the expression of BADPCT from pBAD/gIII(PCT1) in strain TOP10F' failed and no recombinant proteins were detected from whole cells on Coomassie blue-stained SDS-PAGE. Final arabinose concentrations of 0.0002%, 0.002%, 0.02% and 0.2% were used, as recommended by the manufacturer (Invitrogen). Figure 32a shows the lack of recombinant protein from pBAD/gIII(PCT1) in TOP10F'. The plasmid was then transformed into *E.coli* strain LMG194 and the same optimisation conditions used. The results of this optimisation can be seen in Figure 32b. The predicted Mw of the recombinant protein is 37 KDa. Recombinant protein appeared to be present at apparent Molecular weights of 43 KDa and also at 29 KDa (see arrows in Figure 32b). This discrepancy may be due to the high proline content of the linker region, which is thought to be responsible for the anomalous migration of P.69. For comparison, purified His_6-tagged PCT of predicted and apparent molecular weight 30.6 KDa (section 3.2.3.2) was run in parallel (lane 1). Although it is not apparent from Figure 32b, the optimum arabinose concentration for recombinant protein expression was 0.02% (lane 4) and was therefore used in further experiments. The proteins produced were named BADPCTa (43 KDa) and BADPCTb (29 KDa).

Figure32a: SDS-PAGE of whole cell lysates to determine expression of BADPCT from *E. coli* strain TOP10F' pBAD/gIII(PCT1) under the conditions used.

KDa: 10KDa protein ladder

Lane 1: Purified PCT (section 3.2.3.2)

Lane2: Not induced

Lane3: Induced with 0.0002% L-arabinose

Lane4: Induced with 0.002% L-arabinose

Lane5: Induced with 0.02% L-arabinose

Lane6: Induced with 0.2% L-arabinose

Figure32b: SDS-PAGE of whole cell lysates to show the optimisation of BADPCT expression from *E. coli* strain LMG194 pBAD/gIII(PCT1) with various arabinose concentrations. Arrows show protein bands at 43 KDa and 29 KDa.

KDa: 10KDa protein ladder

Lane 1: Purified PCT

Lane2: Not induced

Lane3: Induced with 0.0002% L-arabinose

Lane4: Induced with 0.002% L-arabinose

Lane5: Induced with 0.02% L-arabinose

Figure 32a

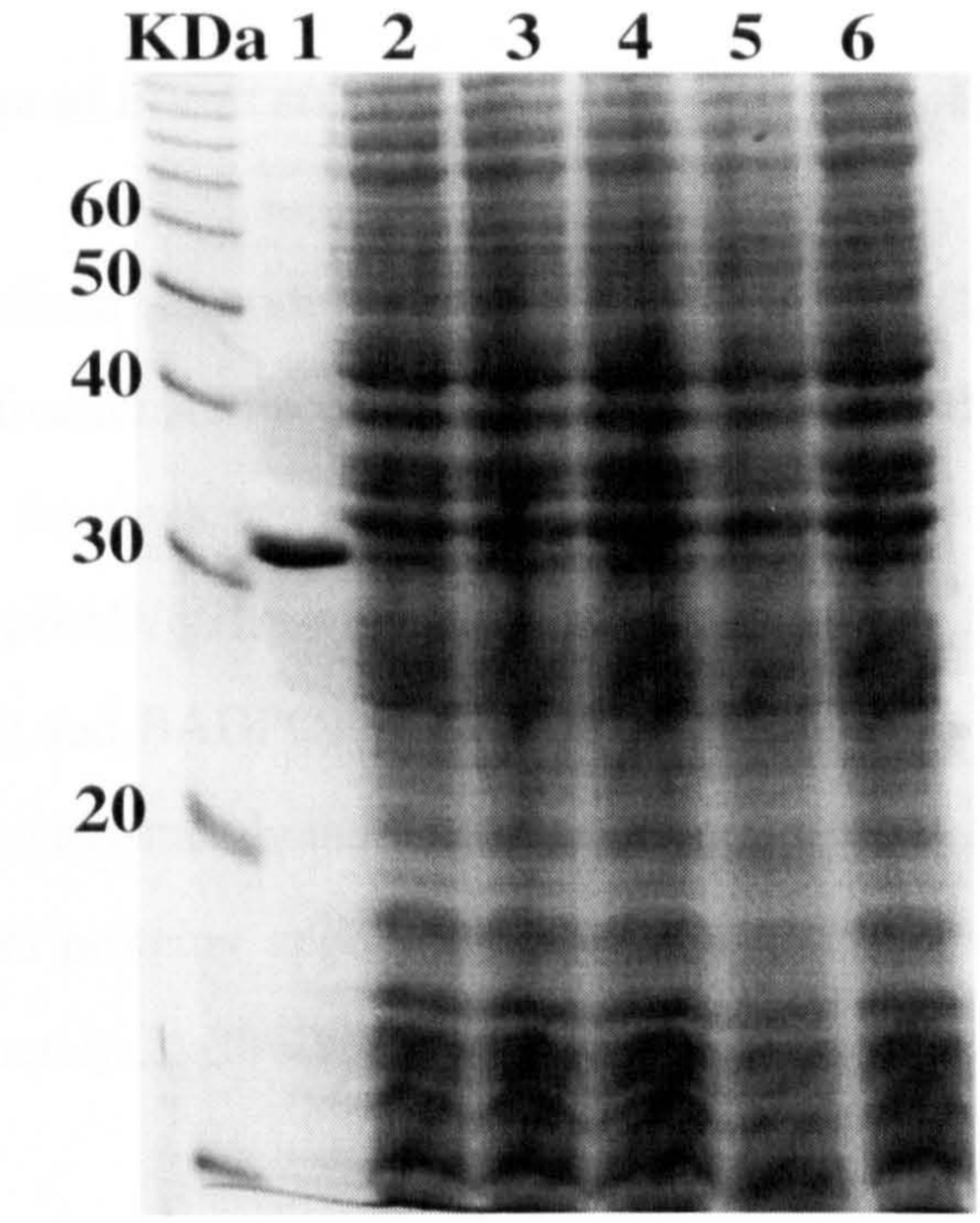
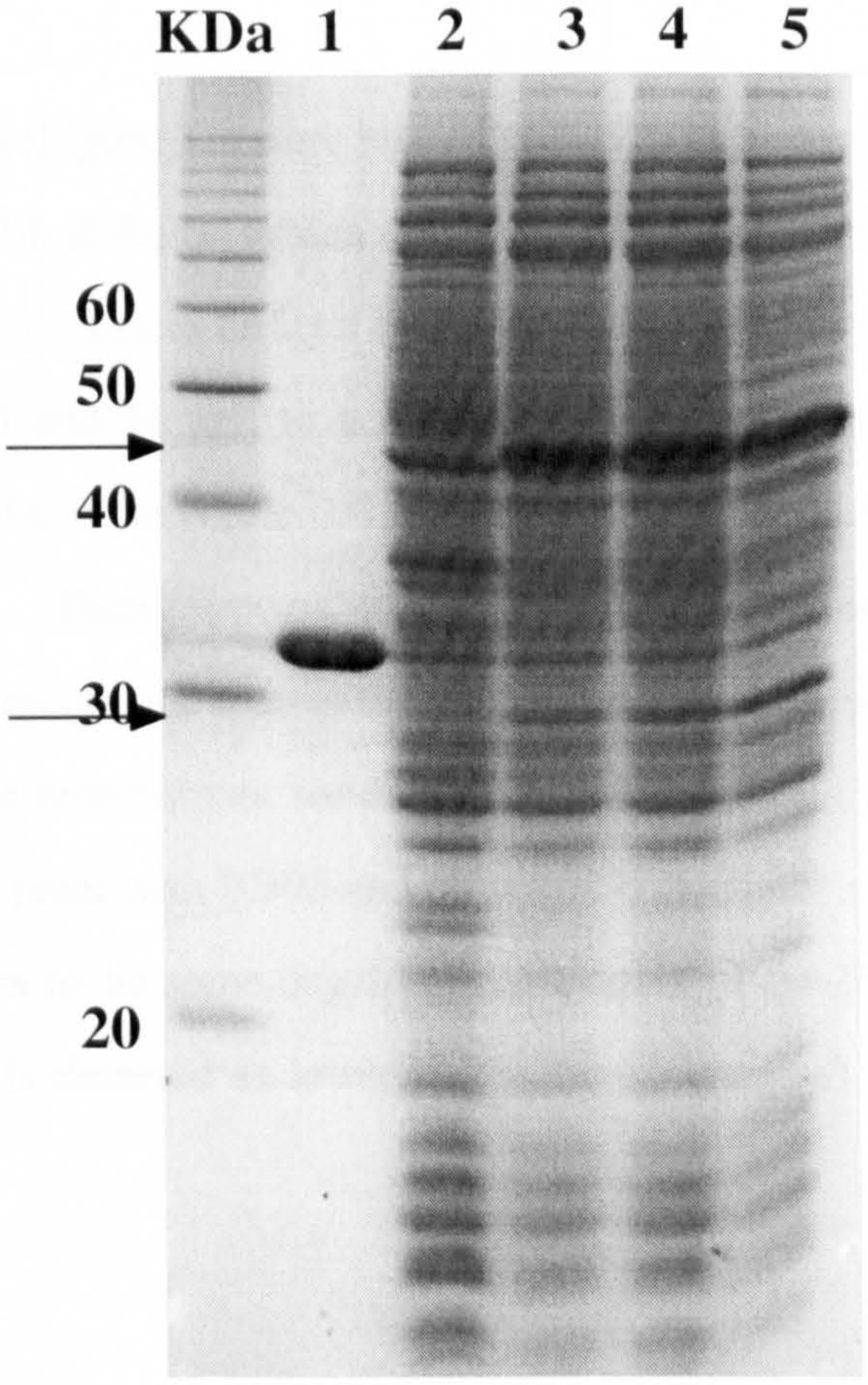


Figure 32b



3.3.3 Localisation of BADPCT

3.3.3.1 Fractionation of *E. coli* strain LMG194 containing pBAD/gIIIA(PCT1)

To determine the cellular location of the BADPCTa and BADPCTb recombinant proteins, cells were fractionated according to section 2.14.1 and the fractions analysed by SDS-PAGE. All fractions can be seen in Figure 33. As a negative control, strain LMG194 containing pBAD/gIIIA with no insert was induced and fractionated. Proteins BADPCTa (43 KDa) and BADPCTb (29 KDa) can be seen in the whole cell fraction, inclusion bodies and outer-membrane fractions (lanes 2, 4 and 10 respectively, Figure 33). The recombinant proteins appear to constitute a major component of the outer membrane (approximately 25-30%).

3.3.3.2 Immunoblot of LMG194 containing pBAD/gIIIA(PCT1) fractions

The cellular fractions (Figure 33) were blotted with a mouse monoclonal antibody BB05 (kindly provided by M. Roberts) which reacts with the linker region of pertactin (which is rich in PQP repeats). As can be seen in Figure 34, the antibody reacted with proteins present in lanes 2, 4 and 10 and to a much lesser extent, lane 8. This confirms the presence of BADPCTa in the whole cell fraction, inclusion bodies, outer membrane and the inner membrane. There appears to be less recombinant protein in the inclusion bodies and the inner membrane than in the outer membrane. The main reacting band is at 43 KDa. The smaller recombinant band of an apparent Mw of 29 KDa in SDS-PAGE (Figure 33) does not react with BB05 and presumably does not contain the linker region target. There appears to be some degradation products of the linker-containing protein (BADPCTa) which is detected as lower molecular weight (<43 KDa) bands in Figure 34.

Figure 33: SDS-PAGE to show fractionation of *E. coli* strain LMG194 containing pBAD/gIII(PCT1) or pBAD/gIID(no insert) after induction with 0.02% arabinose. Arrows identify bands at 43 KDa and 29 KDa.

KDa: 10 KDa protein ladder

Lane 1: Whole-cell lysate pBAD/gIII(no insert)

Lane2: Whole-cell lysate pBAD/gIII(PCT1)

Lane3: Inclusion-body fraction pBAD/gIII(no insert)

Lane4: Inclusion-body fraction pBAD/gIII(PCT1)

Lane5: Soluble fraction (cytoplasm and periplasm) pBAD/gIII(no insert)

Lane6: Soluble fraction (cytoplasm and periplasm) pBAD/gIII(PCT1)

Lane 7: Inner-membrane fraction pBAD/gIII(no insert)

Lane 8: Inner-membrane fraction pBAD/gIII(PCT1)

Lane 9: Outer-membrane fraction pBAD/gIII(no insert)

Lane 10: Outer-membrane fraction pBAD/gIII(PCT1)

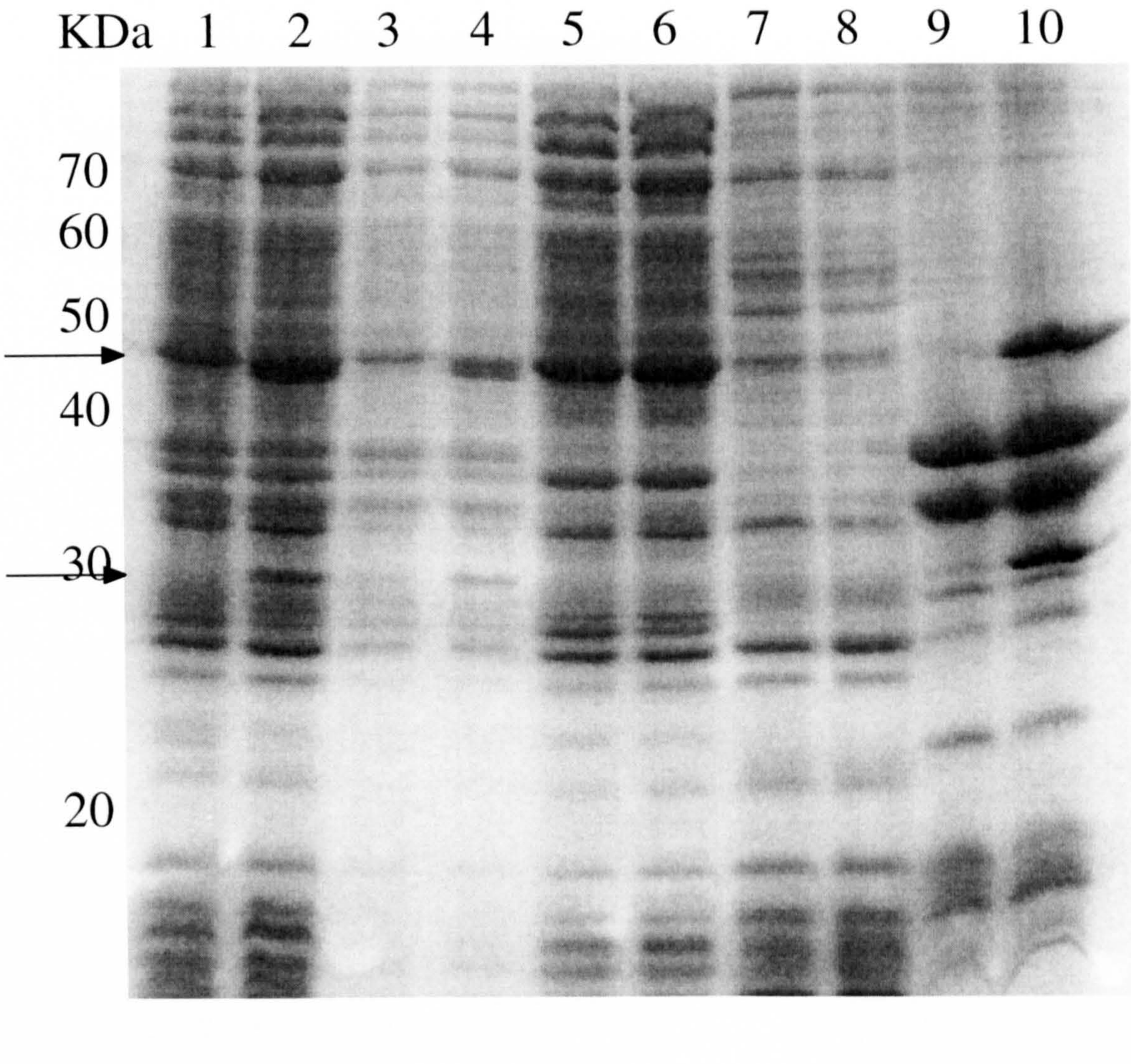

KDa
1
2
3
4
5
6
7
8
9
10
70
60
50
40
30
20

Figure 34: Immunoblot with monoclonal antibody BB05 of the samples described in figure 33

KDa: 10 KDa protein ladder

Lane 1: Whole-cell lysate pBAD/gIII(no insert)

Lane2: Whole-cell lysate pBAD/gIII(PCT1)

Lane3: Inclusion-body fraction pBAD/gIII(no insert)

Lane4: Inclusion-body fraction pBAD/gIII(PCT1)

Lane5: Soluble fraction (cytoplasm and periplasm) pBAD/gIII(no insert)

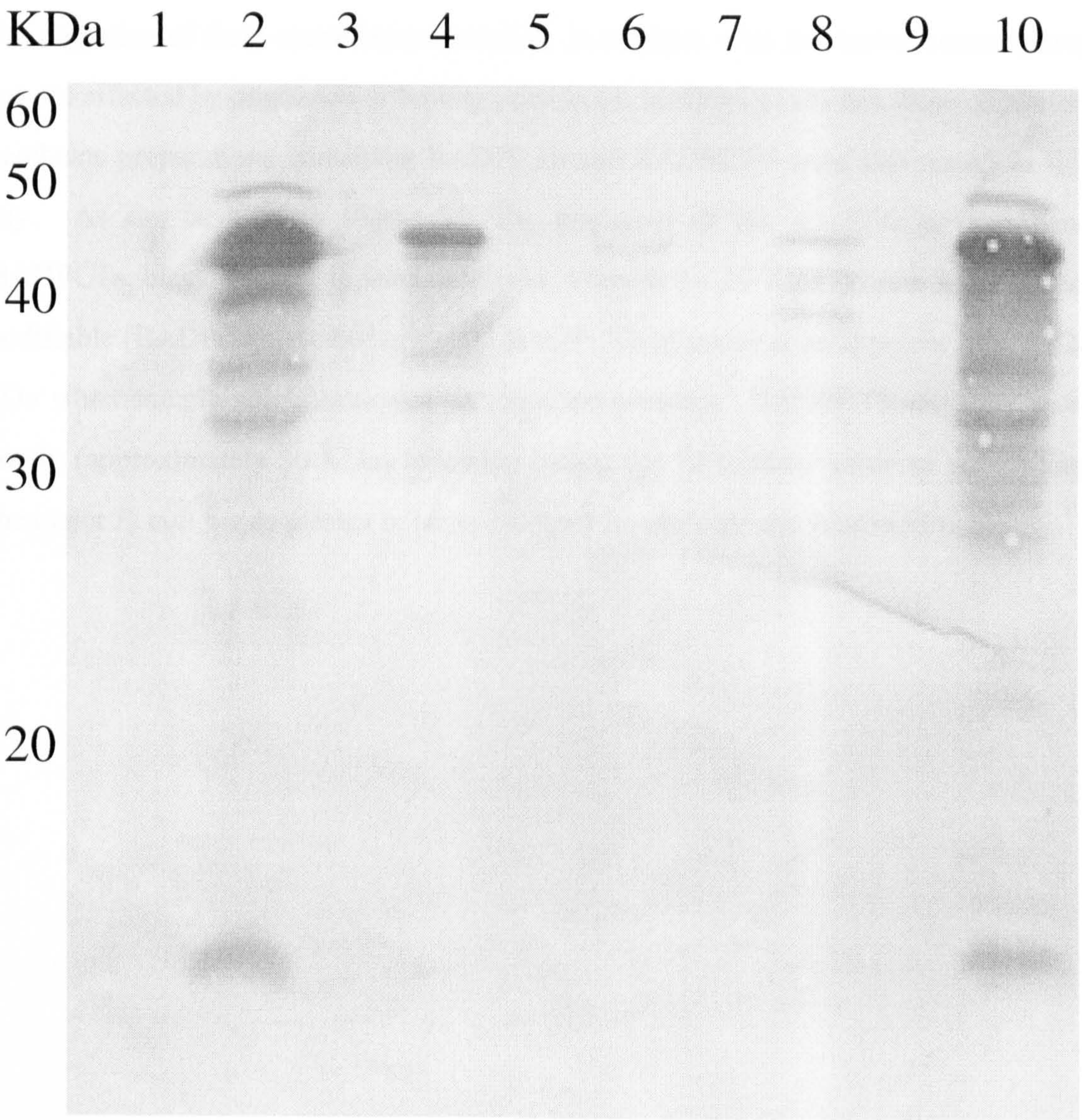
Lane6: Soluble fraction (cytoplasm and periplasm) pBAD/gIII(PCT1)

Lane 7: Inner-membrane fraction pBAD/gIII(no insert)

Lane 8: Inner-membrane fraction pBAD/gIII(PCT1)

Lane 9: Outer-membrane fraction pBAD/gIII(no insert)

Lane 10: Outer-membrane fraction pBAD/gIII(PCT1)


KDa
1
2
3
4
5
6
7
8
9
10
60
50
40
30
20

3.3.4 Heat modifiability of BADPCTa and BADPCTb

The native C-terminal domains of BrkA and Tcf are heat modifiable *ie.* the migration of these proteins on SDS-PAGE depends on the temperature of solubilisation of the sample (R. Parton, personal communication). To assesss the heat modifiability of the recombinant C-terminal domains, inclusion bodies were resuspended in PBS and either boiled for 10 min or were maintained at room temperature prior to loading onto the gel. The migration of these recombinant proteins, in contrast with the native counterparts, was not affected by ommision of boiling prior to gel loading (results not shown). Outer-membrane preparations containing BADPCTa and BADPCTb were also tested in this way. As can be seen in Figure 35, the migration of the 43 KDa protein band (BADPCTa, black arrows) appears unaltered, whereas the 29 KDa protein band is heat modifiable (BADPCTb, dashed arrows). BADPCTb migrates at an apparent Mw of 26 KDa when sample solubilisation is at room temperature. BADPCTb migrates more slowly (approximately 30 KDa) following boiling for 10 minutes prior to gel loading. The major *E. coli* bands present in lanes 3-6 (grey arrows) are also heat modifiable.

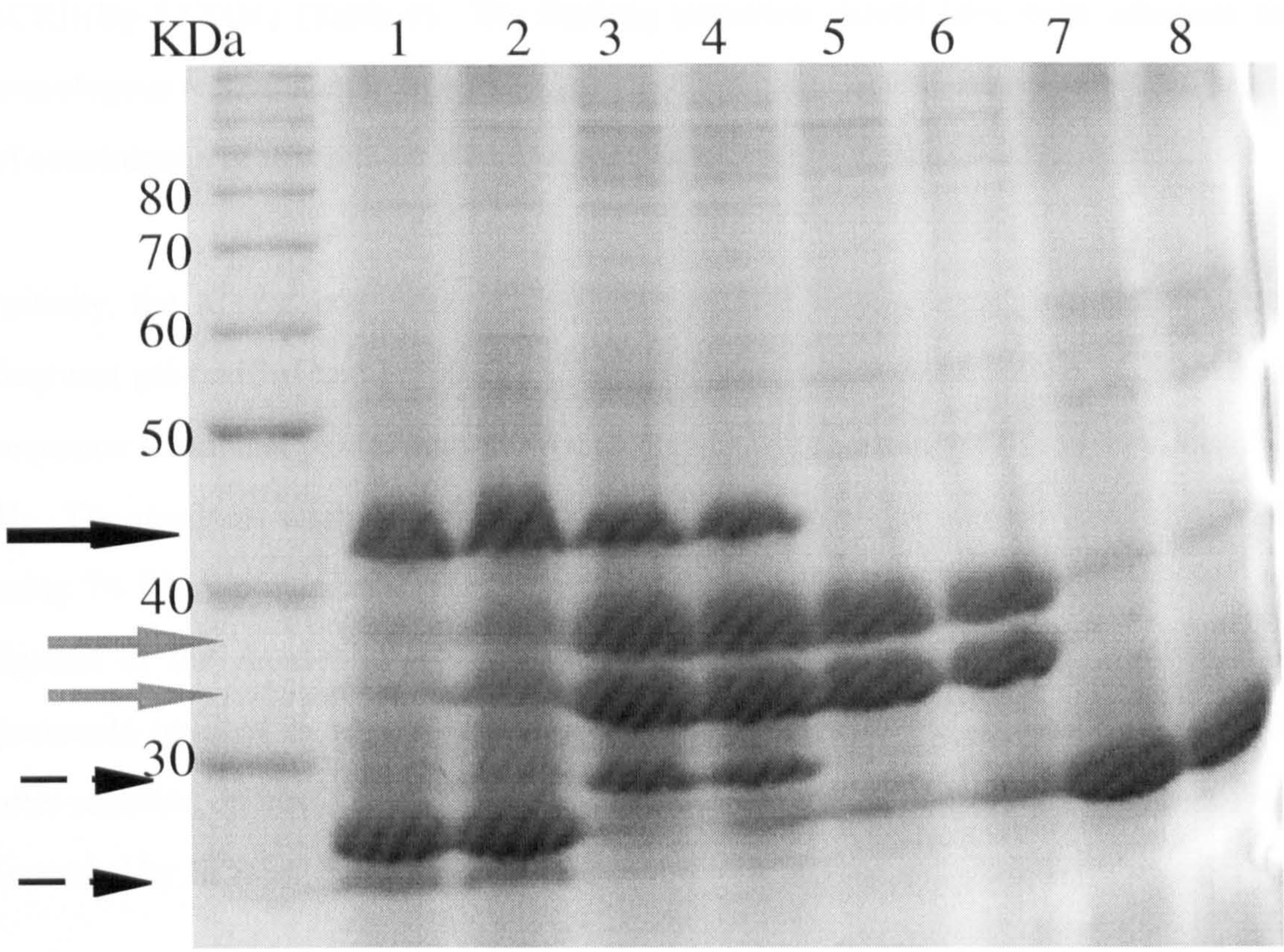
Figure 35: Heat modifiability of BADPCTa and BADPCTb (black arrows and dashed arrows respectively). *E. coli* proteins are marked with grey arrows.

Lanes 1 and 2: Outer-membrane fraction of *E. coli* containing pBAD/gIII(PCT1) without boiling

Lanes 3 and 4: Outer-membrane fraction of *E. coli* containing pBAD/gIII(PCT1) with boiling

Lanes 5 and 6: Outer-membrane fraction of *E. coli* containing pBAD/gIII(no insert) with boiling

Lanes 7 and 8: Outer-membrane fraction of *E. coli* containing pBAD/gIII(no insert) without boiling


KDa
1
2
3
4
5
6
7
8
80
70
60
50
40
30

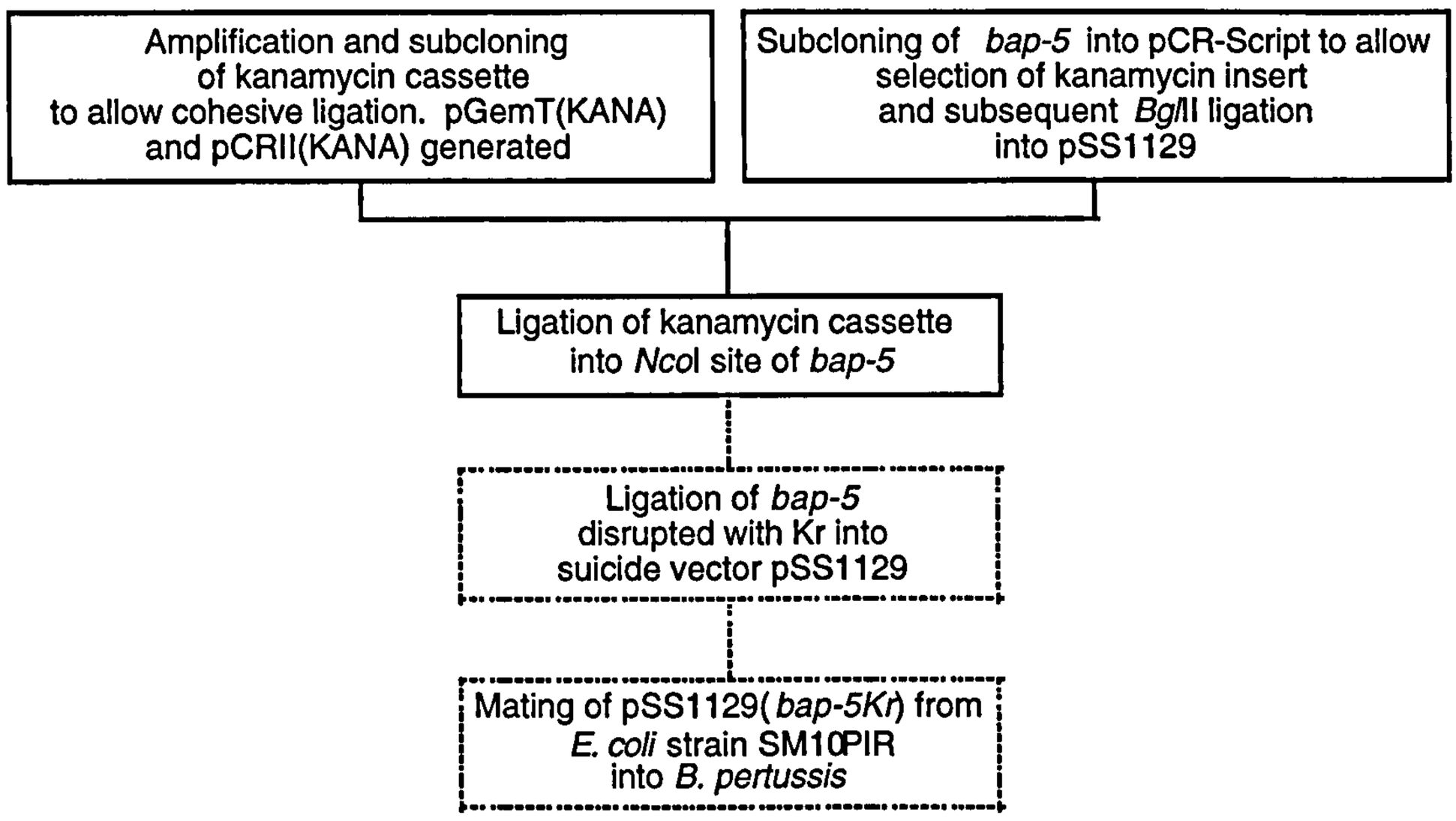
3.4 Construction of a *bap-5* mutant of *B. pertussis*

The intention was to disrupt the *bap-5* open reading frame with a kanamycin cassette (K^R). This mutant would be of great value in assessing the function of the Bap-5, perhaps as a virulence factor in a number of *in vivo* and *in vitro* assays such as mouse respiratory tract colonisation and adhesion. The *Nco*I site within *Bap-5* was considered to be a suitable site for insertion of K^R as such a disruption would provide flanking sequences of 1738 and 542 bp and only one *Nco*I site was present in the vector pCRII(Bp-5ATG1) (Table 4). The flanking sequences should have been adequate for homologous recombination into the *B. pertussis* genome. Unfortunately, the final stages of construction of this mutant were not completed, due to technical difficulties.

Initially, the K^R cassette from pUC4K was excised using *Bam*HI and the resultant fragment gel-purified and polished. The presence of an *Nco*I site in the pCRII vector sequence meant that partial digests were necessary, with serial dilutions of *Nco*I used for 1h. The resultant single digest products (Figure 37) were gel-purified and made blunt using T4 DNA polymerase. Attempts to disrupt the cloned *bap-5* gene with K^R by ligation of the *Bam*HI K^R fragment at the *Nco*I using standard and rapid ligation protocols resulted in no colonies following transformation into *E. coli* strain JM109 after selection on media containing ampicillin (encoded by pCRII vector) and kanamycin (encoded by K^R).

An alternative strategy was attempted which would allow cohesive ligation of K^R into *bap-5*. This is outlined in Figure 36. The kanamycin cassette was amplified from pUC4K using primer KANANCOI (Table 5) which contained an *Nco*I site and annealed to 2 sites which flank the kanamycin cassette. The resultant amplicon was cloned into pGemT and pCRII-TOPO to generate plasmids pGemT(KANA) and pCRII(KANA) respectively. Such plasmids will enable the cloning of the K^R cassette into a greater variety of restriction sites and should be a useful laboratory resource.

Figure 36: Construction of a Bap-5 mutant of *B. pertussis*[*]



*Note: Dashed lines respresent the uncompleted stages of mutant construction

To enable selection for kanamycin resistance within the disrupted *bap-5* gene, *bap-5* was cloned into pCR-Script following amplification with primers BAP5MUT5' and BAP5MUT3' (Table 5). These primers possessed *Bgl*II restriction sites to enable subsequent ligation of *bap-5K*^{R} into the *Bgl*II site of the suicide vector pSS1129. The PCR was performed using the Expand™ high fidelity PCR system (section 2.7.3) with pCRII(ATG1) as a template. The resultant amplicon (approximately 3.7 Kb) was gel purified and ligated into pCR-Script (section 2.9.1.2) which contains no *Nco*I sites and is not kanamycin resistant. Recombinant clones were selected using LB agar containing both ampicillin and kanamycin and were checked by *Nco*I analysis. One positive clone was selected and named pCR-Script(ATG1).

The cloned K^R cassette in pCRII(KANA) was digested with *Nco*I and the 1.3 Kb insert (Figure 37a) was gel purified, ligated into the *Nco*I site of pCR-Script(ATG1) using the rapid ligation protocol (section 2.9.2.3) and transformed into *E. coli* strain JM109. Resultant colonies were checked for presence of a correct size insert by *Nco*I restriction analysis of plasmids. The positive selection (kanamycin) resulted only in clones that contained the kanamycin cassette inserted into *bap-5*. Figure 37b confirms the presence of the 1.3 Kb K^R cassette. One clone was chosen and called pCR-Script(Bap-5 K^R). Insertion of the kanamycin cassette at the correct site was confirmed by sequencing using primers PEB1FOR12 and PEB1REV6 (Table 5).

The plasmid pCR-Script (Bap-5K^R), containing *bap-5* disrupted with K^R was digested with *Bgl*II and the disrupted *bap-5* was gel purified. The suicide vector pSS1129 was also digested with *Bgl*II and gel purified. Attempts were made to ligate the *bap-5K*^{R} into pSS1129 using both standard ligation and rapid ligation procedures. The reactions were transformed into *E. coli* strain TOP10F' by heat shock, or *E. coli* strain JM109 by electroporation. Insert : vector ratios of 1:1, 3:1, 10:1 and 15:1 were used. Also, the vector was dephosphorylated in an attempt to increase the chance of insert incorporation. In addition, an insert-free control was used to check the ability of the vector to self ligate. Selection was performed on LB agar containing ampicillin and

kanamycin or LB agar containing ampicillin alone (insert-free control) and resulted in either no colonies on any plates or, in some instances, few colonies after 2 days which later proved to contain a plasmid with properties different to those expected/predicted. Despite several attempts, ligation of the disrupted *bap-5* gene into pSS1129 was not successful and for unknown reasons it was not possible to regenerate pSS1129 by self ligation following *Bgl*II digestion. If the above cloning experiment had worked, the suicide construct could then have been mated into *B. pertussis* and construction of the mutant completed.

Figure 37a: Digestion of the pGemT(KANA) (lanes 1 and 2) and PCRII (KANA) (lanes 3 - 6) with *Nco*I to show release of a 1.3 Kb kanamycin resistance cassette (K^R) (arrows). The other bands are vector DNA. pGemT has an *Nco*I site which is also cut (lanes 1 and 2).

Figure 37b: Digestion of pCR-Script(Bp-5 K^r) with *Nco*I to show the 1.3 Kb K^R is present (arrow against lane 1). The 5.5 Kb band is the vector and *bap-5* sequences and the 7.5 Kb band is presumably a partially cut or uncut form of pCR-Script(Bp-5K^r).

Figure 37a

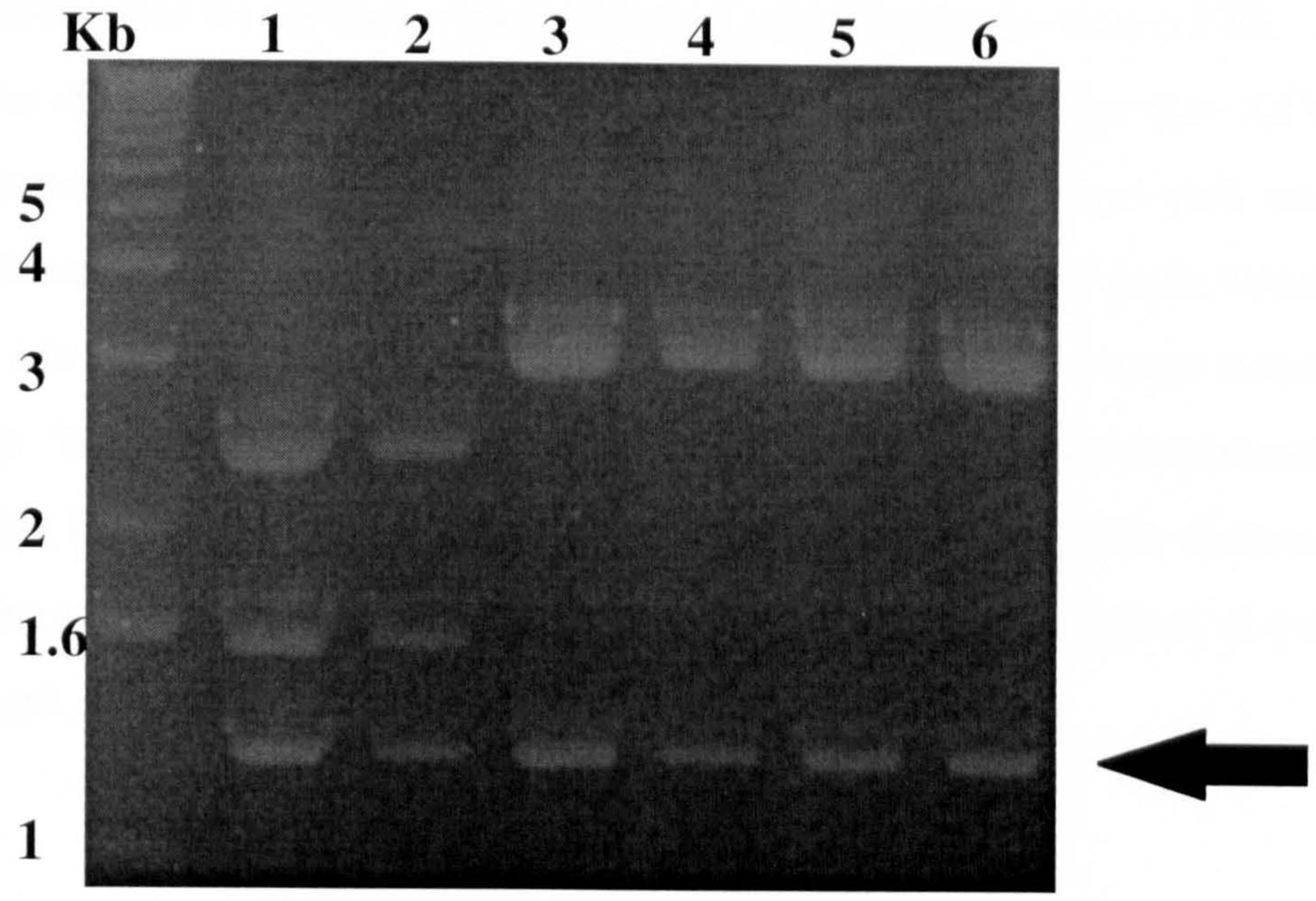
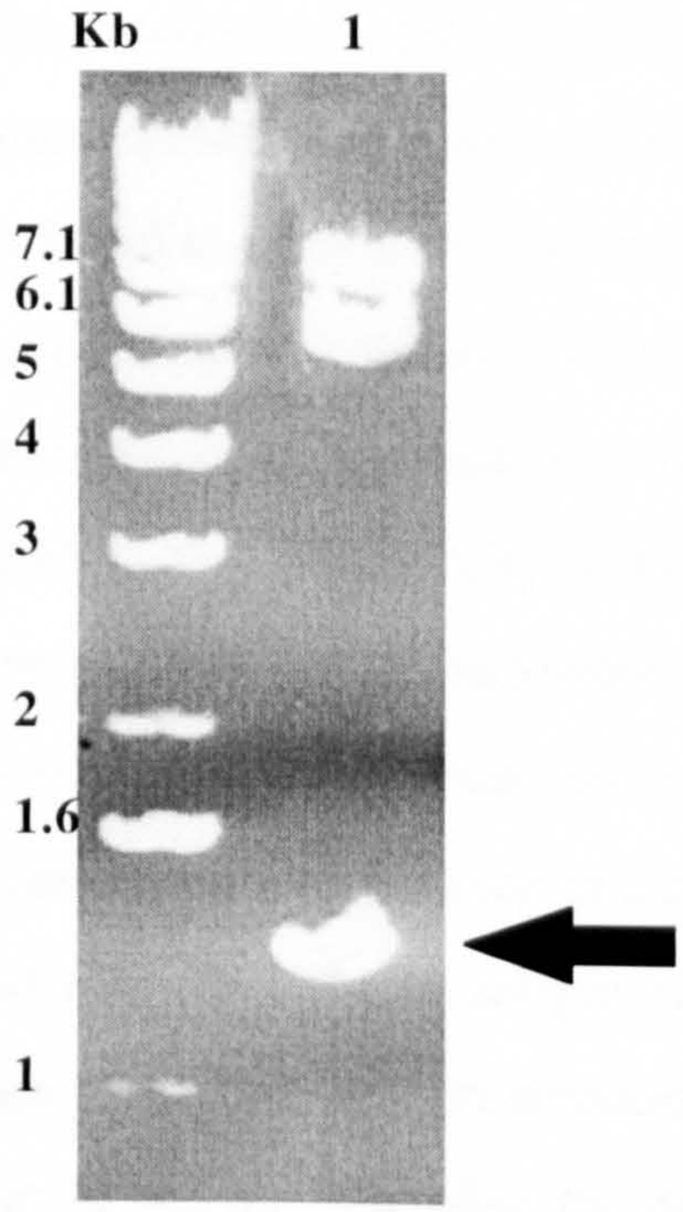


Figure 37b



3.5 Mouse protection tests using autotransporter domains

The methods used for the mouse protection tests are described in section 2.18. Briefly, two doses of vaccine (15µg of total protein per dose or 1/10th human dose ACV) were administered subcutaneously with a two week interval, followed one week later by a sublethal intranasal challenge (of $1x10^6$ cfu/mouse) with *B. pertussis* strain Taberman or 18323. Mouse lungs were removed seven days post challenge and viable counts were obtained. The lung count data can be found in Appendix III. Three experiments were performed, R12, R13 and R14, and the results from mice given acellular vaccine (ACV) containing pertactin (P69) or alhydrogel (AL) were combined and analysed in a separate experiment.

3.5.1 Mouse protection test R12

This experiment was designed to test the ability of Bap-5, and NTS (Bap-5 specific portion), as urea-solubilised inclusion bodies, to protect adult mice against challenge with either of two *B. pertussis* strains (Taberman or 18-323). Strain Taberman was used as this strain was the first found to contain the *bap-5* gene (section 3.1). Strain 18-323 was also used as this is a well-described, mouse-virulent strain. As a positive control in this experiment, a Smithkline Beecham acellular vaccine (ACV) containing P.69, FHA and PTd was used at 1/10th human dose. Also included as a positive control was a purified pertactin (P.69) preparation, also from Smithkline Beecham. Negative controls included alhydrogel (AL) alone and *E. coli* crude cell lysate. As can be seen in Figure 38a and Table 7a, both challenge strains behaved in a similar way, in terms of the lung counts of mice immunised with the different antigen preparations. No interaction was seen between the challenge strains and antigen preparations (confirmed by analysis of variance [ANOVA], Table 7a and b). Therefore the data for each challenge strain was combined to increase the sample size and analysed as one experiment. Figure 38b and Table 7B2 gives the results of such analysis.

Figure 38a: Results of Experiment R12: Total lung counts (log_{10}) from mice immunised with various *B. pertussis* antigen preparations and challenged intranasally with two different strains (Taberman and 18-323) of live *B. pertussis*. Each point represents one mouse. Counts of <100 or >$3x10^6$ were given notional values of 20 and 10^7 respectively and plotted as open symbols. The Minitab 'jitter' option was applied so as to spread the points horizontally to avoid coincidence.

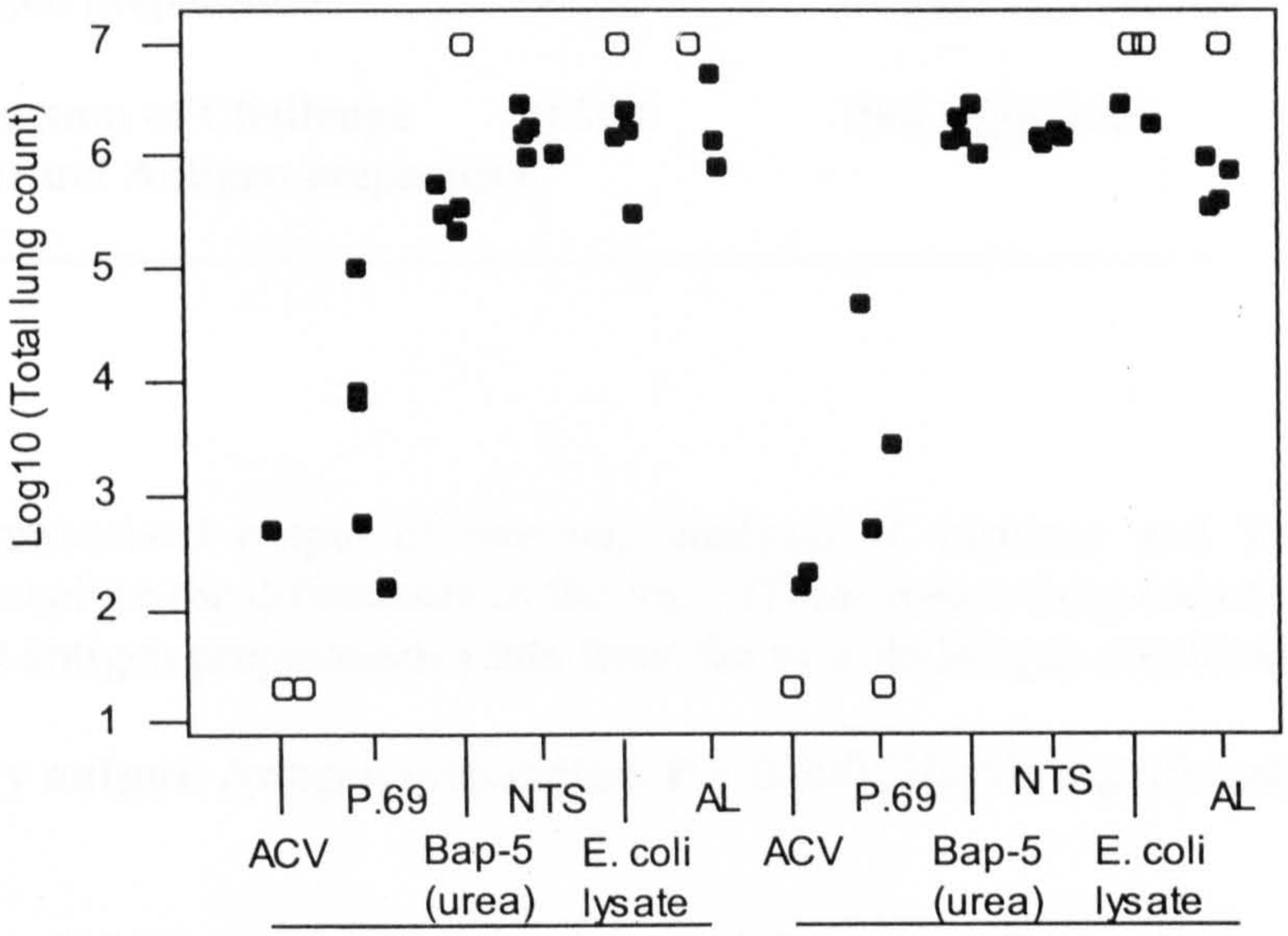


Figure38b: Boxplots of the data in Figure 37a (above) with the results of the two challenges combined.

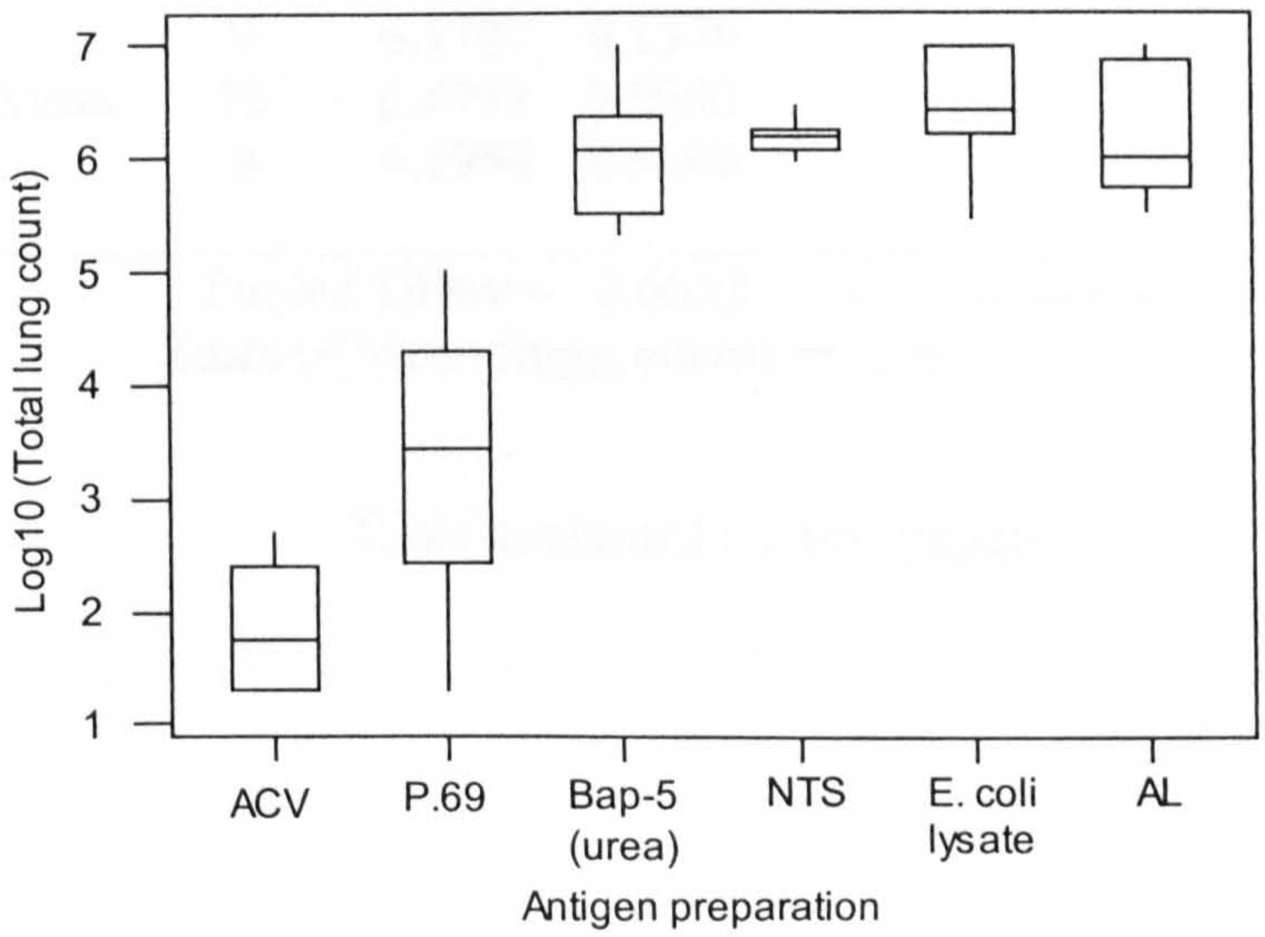


Table 7: Expt. R12: Summary of the statistical analysis of the data in Fig 38.

Part A. Summarised output from Minitab of the *two-way* analysis of variance to examine the influence of *Antigen preparation* and *Challenge strain* on the *log_{10} (Total mouse lung count)*.

Source of variation	P-value	Interpretation
Challenge strain	0.915	Not significant
Antigen preparation	0.000	Highly significant
Interaction of Challenge strain and Antigen preparation	0.500	Not significant

Part B. Summarised output of *one-way* analysis of variance and Tukey test, from Minitab, to explore the differences in the *log_{10} (Total mouse lung count)* associated with the different antigen preparations (data from the two challenges combined).

B1. Primary output: Antigen preparation: P = 0.000 (Highly significant)

B2. Basic statistics and graphical summary (Mean and 95% confidence intervals)

Antigen Preparation	No. obs.	Mean	StDev	Mean log_{10} count* (and 95% CI)
ACV	6	1.8465	0.6220	(---*--)
P.69	9	3.3096	1.1963	(--*-)
Bap-5 (urea)	10	6.0220	0.5128	(--*-)
NTS	9	6.1797	0.1379	(--*-)
E. coli lysate	10	6.4999	0.5050	(--*-)
AL	9	6.1959	0.5699	(--*--)
	Pooled StDev =	0.6652		--+---------+---------+---------+----
	Scale of Mean (log_{10} count) →			1.6 3.2 4.8 6.4

(Table continued on next page)

Table 7. (Continued)

B3. Tukey's pairwise comparisons of lung-count means (log_{10} values)
The table shows the 95% confidence intervals for {*Column antigen mean* minus *Row antigen mean*}. Significant differences in counts are shown in **bold.**

	ACV	P.69	Bap-5 (urea)	NTS	*E.coli* lysate
P.69	**-2.5042** **-0.4219**				
Bap-5 (urea)	**-5.1956** **-3.1554**	**-3.6201** **-1.8047**			
NTS	**-5.3744** **-3.2921**	**-3.8014** **-1.9389**	-1.0654 0.7499		
E. coli lysate	**-5.6735** **-3.6333**	**-4.0980** **-2.2827**	-1.3614 0.4055	-1.2278 0.5875	
AL	**-5.3905** **-3.3082**	**-3.8175** **-1.9550**	-1.0815 0.7338	-0.9474 0.9151	-0.6036 1.2117

When one way Anova was applied to the combined results, highly significant differences between the antigen preparations were apparent. This was further analysed using the Tukey test which revealed that the results with ACV were significantly different from those with all of the other antigen preparations Table 7B3. Immunisation with ACV resulted in significantly lower lung counts and therefore better protection. P.69 was significantly different from both ACV and the other antigen preparations and also gave significant protection. The Bap-5 (urea) and NTS antigen preparations were ineffective immunogens because the lung counts were not significantly different to those obtained with alhydrogel (AL).

3.5.2 Mouse protection test R13

Only strain Taberman was used for this experiment, on the basis of the results from experiment R12. This experiment involved using the urea-solubilised C-terminal domains of pertactin (PCT), BrkA (BCT), Tcf (TCT) and Bap-5 (Bap5CT). The results were treated as in R12 except for the omission of the challenge differences. ACV was the only preparation to give significant protection in this experiment (Figure 39 and Table 8).

Figure 39a: Results of Experiment R13: Total lung counts (log_{10}) from mice immunised with various *B. pertussis* antigen preparations and challenged intranasally with live *B. pertussis* strain Taberman. Each point represents one mouse. Counts of <100 or >$3x10^6$ were given notional values of 20 and 10^7 respectively and plotted as open symbols. The Minitab 'jitter' option was applied so as to spread the points horizontally to avoid coincidence.

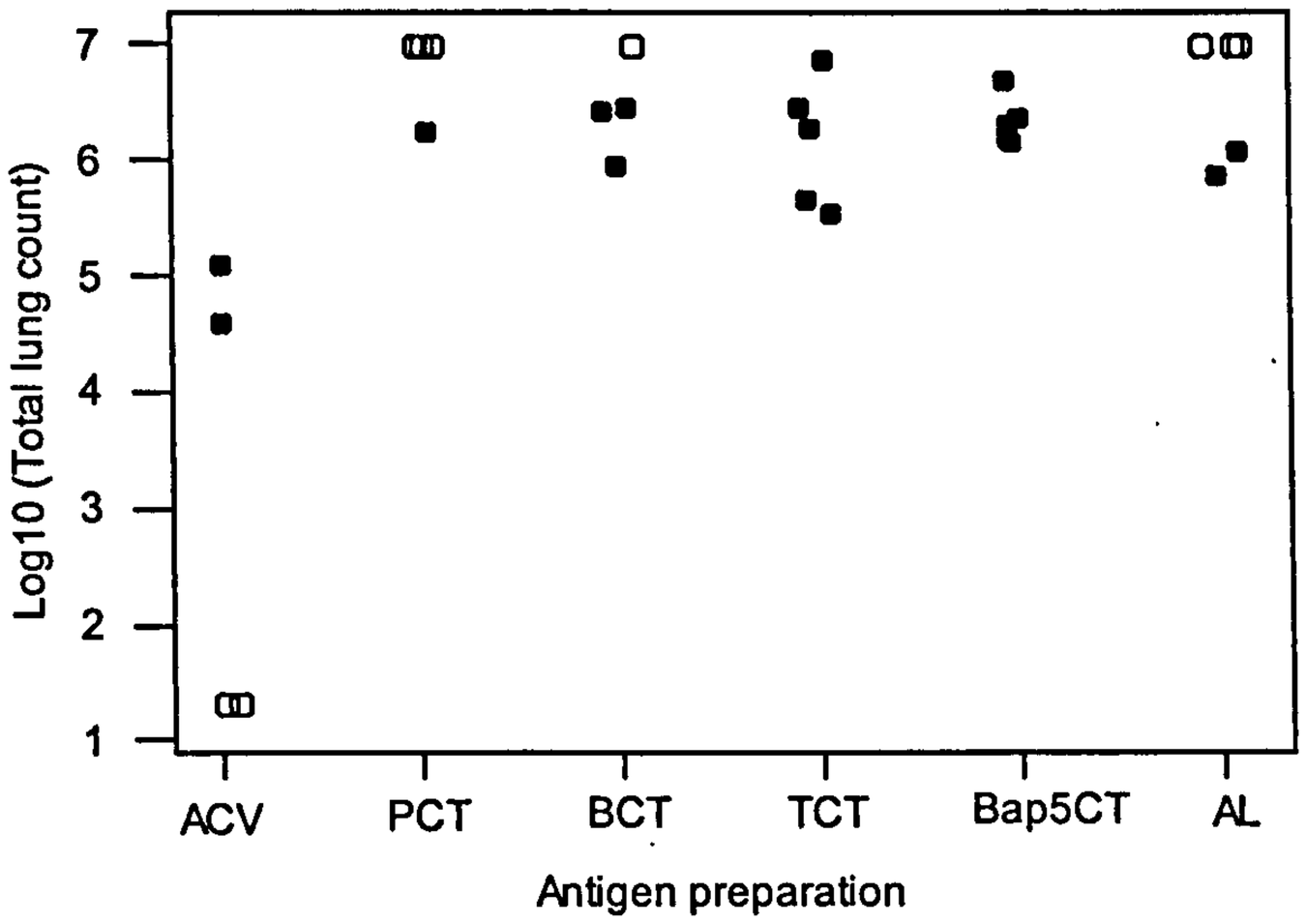


Figure 39b: Boxplots of the above data

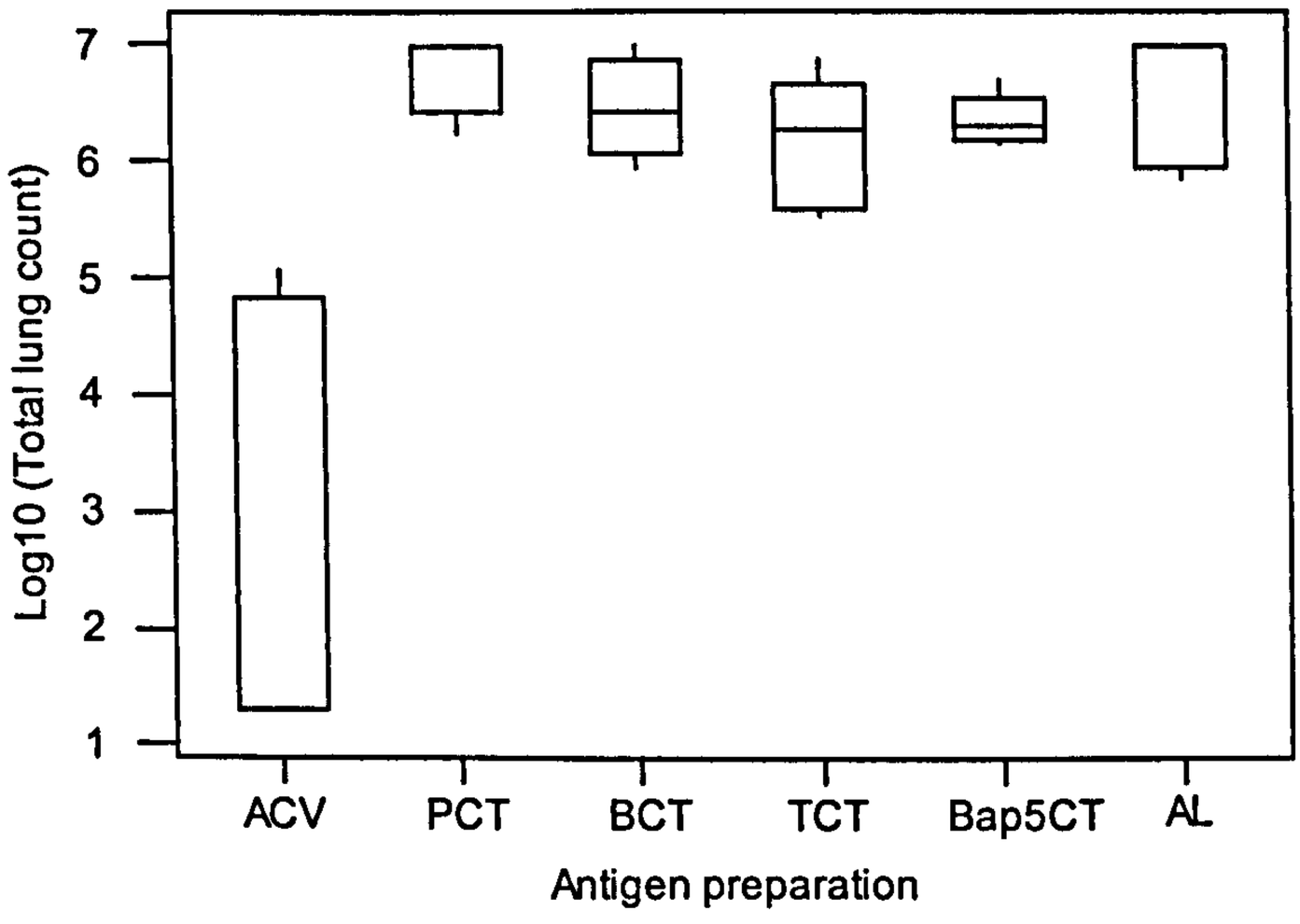


Table 8: Experiment R13: Summary of the statistical analysis of the data in Fig 39. Output of *one-way* analysis of variance and Tukey test, from Minitab, to explore the differences in the *log_{10} (Total mouse lung count)* associated with the different antigen preparations.

1. Primary output of the one-way Anova: Antigen preparation: P = 0.000 (Highly significant)

2. Basic statistics and graphical summary (Mean and 95% confidence intervals)

Antigen Preparation	No. obs.	Mean	StDev	Mean log_{10} count and 95% CI
ACV	5	2.7169	1.9460	(----*----)
PCT	4	6.8138	0.3724	(-----*-----)
BCT	4	6.4582	0.4275	(-----*-----)
TCT	5	6.1556	0.5619	(----*-----)
Bap5CT	5	6.3572	0.2145	(-----*----)
AL	5	6.5873	0.5705	(----*-----)
Pooled StDev =		0.9259		--------+--------+--------+------
Scale of Mean (log_{10} count) ➡				3.2 4.8 6.4

3. Tukey's pairwise comparisons of lung-count means (log_{10} values)
The table shows the 95% confidence intervals for {*Column antigen mean* minus *Row antigen mean*}. Significant differences in counts are shown in **bold**.

	ACV	PCT	BCT	TCT	Bap5CT
PCT	**-6.0338** **-2.1601**				
BCT	**-5.6782** **-1.8044**	-1.6860 2.3973			
TCT	**-5.2648** **-1.6126**	-1.2786 2.5952	-1.6343 2.2395		
Bap5CT	**-5.4665** **-1.8142**	-1.4803 2.3935	-1.8359 2.0379	-2.0278 1.6245	
AL	**-5.6966** **-2.0443**	-1.7104 2.1634	-2.0660 1.8078	-2.2579 1.3944	-2.0562 1.5960

3.5.3 Mouse protection test R14

Again, strain Taberman was used as the challenge in this experiment which was designed to test the effect of urea solubilisation on the protective capacity of Bap-5 and P.69. This was necessary as P.69 was present as a non-denatured protein and may not have been a suitable comparison to urea denatured Bap-5. P.69 was denatured in 8 M urea and Bap-5 was either urea-solubilised as before, solubilised in PBS, or dialysed from urea into PBS. Analysis was performed as above and results showed that the ACV was significantly more protective than other antigen preparations with the exception of P.69 and P.69 (urea). Urea treatment of P.69 did not appear to alter its ability to protect against challenge and no significant difference was detected between the Bap-5 preparations. The results of this analysis can be seen in Figure 40 and Table 9.

3.5.4 Combination and analysis of data obtained with acellular vaccine, pertactin and alhydrogel

Attention was directed to the ACV, P.69 and AL which had been tested in experiments R12, R13 and R14. First each antigen was subjected to Anova to determine whether there were significant differences between the results obtained from each experiment (Table 10a). None was detected. Therefore, for each of these three antigen preparations, the results of the three experiments were pooled to allow antigen comparisons based on all the data. The point plots and boxplots (Figures 41a and 41b) suggested that protection from ACV was not significantly more effective than P.69 and both gave significantly better protection than AL based on the lung count data. This was confirmed by analysis of variance and the Tukey test (Tables 10B, 10B2 and 10B3).

Figure 40a: Results of Experiment R14: Total lung counts (log_{10}) from mice immunized with various *B. pertussis* antigen preparations and challenged intranasally with live *B. pertussis* strain Taberman. Each point represents one mouse. Counts of <100 were given notional values of 20 and plotted as open symbols. The Minitab 'jitter' option was applied so as to spread the points horizontally to avoid coincidence.

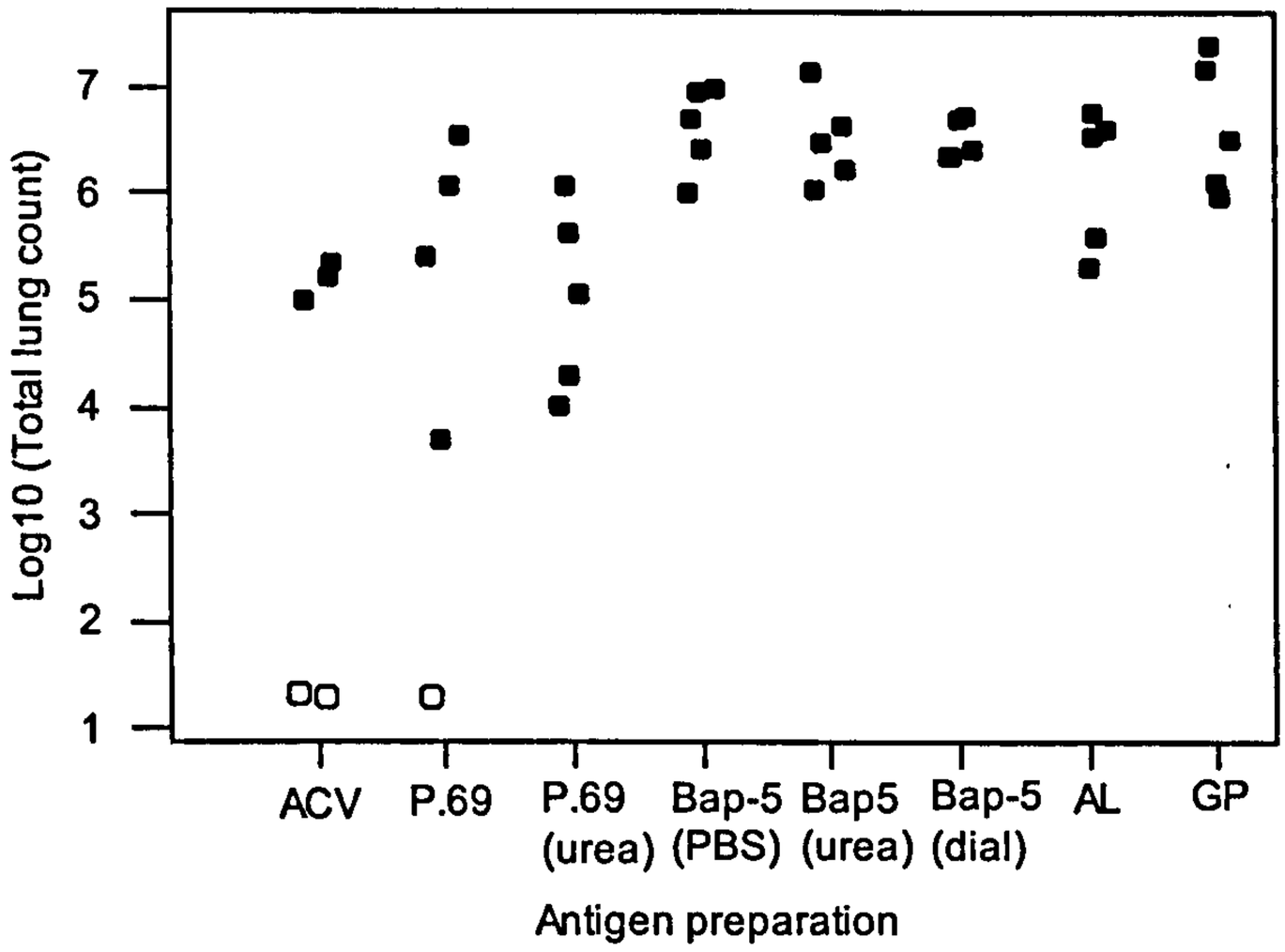


Figure 40b: Boxplots of the above data

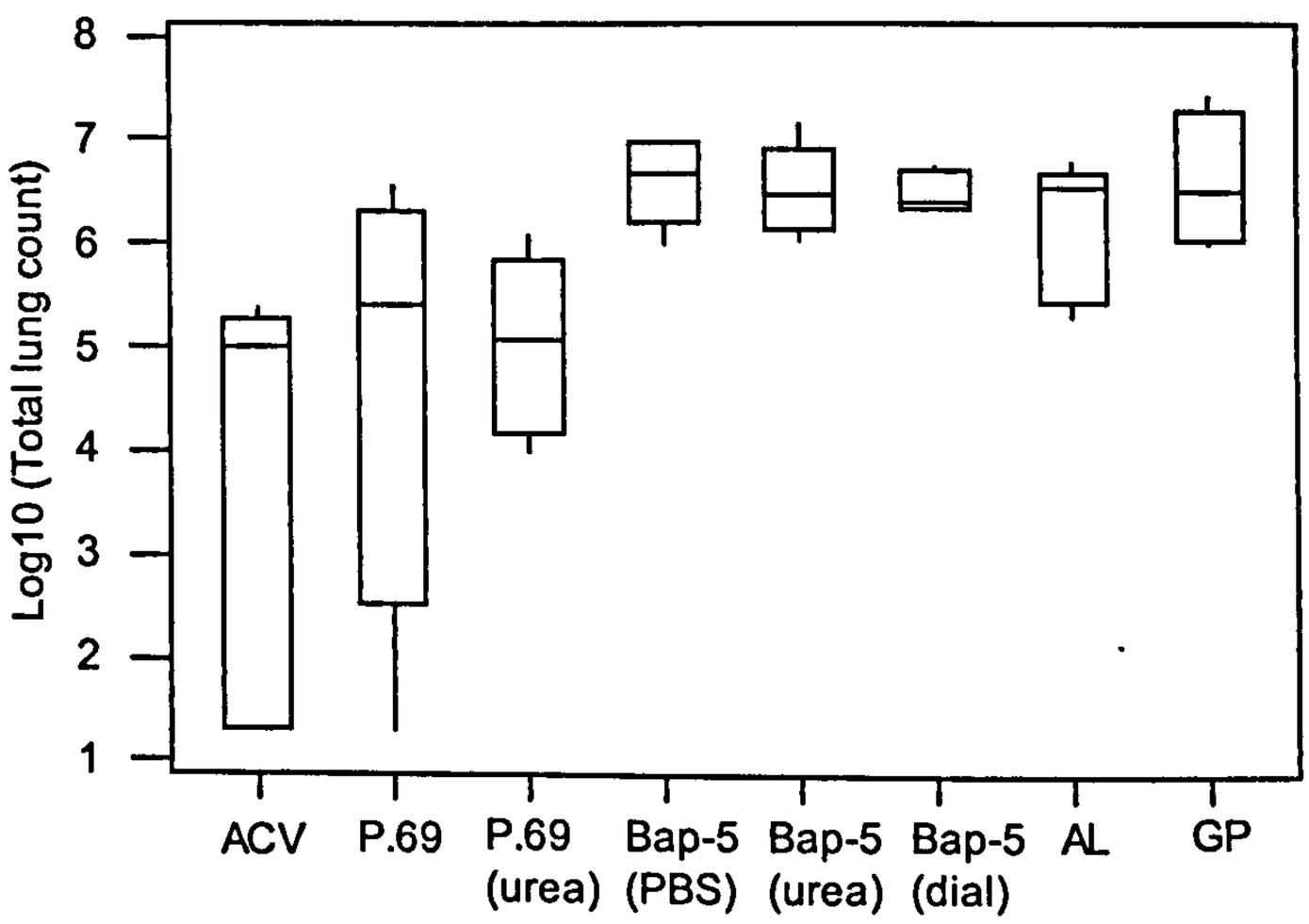


Table 9: Experiment R14. Summarised output of *one-way* analysis of variance and Tukey test, from Minitab, to explore the differences in the *Log_{10} (Total mouse lung count)* associated with the different antigen preparations.

1. Primary output of the one-way Anova: Antigen preparation: P = 0.001 (Highly significant)
2. Basic statistics and graphical summary (Mean and 95% confidence intervals)

Antigen preparation	No. obs.	Mean	StDev	-Mean log_{10} count and 95% CI
ACV	5	3.634	2.133	(------*------)
P.69	5	4.612	2.146	(------*------)
P.69 (urea)	5	5.021	0.876	(------*------)
Bap-5 (PBS)	5	6.617	0.413	(------*------)
Bap-5 (urea)	5	6.520	0.435	(------*------)
Bap-5 (dial)	5	6.522	0.190	(------*------)
AL	5	6.170	0.669	(------*------)
GP	5	6.650	0.633	(------*------)
			Pooled StDev = 1.182	---+---------+---------+---------+---------+---
			Scale of Mean (log_{10} count)	➡3.0 4.5 6.0 7.5

3. Tukey's pairwise comparisons of lung-count means (log_{10} values)
The table shows the 95% confidence intervals for {*Column antigen mean* minus *Row antigen mean*}. Significant differences in counts are shown in **bold.**

	ACV	P.69	P.69 (urea)	Bap-5 (PBS)	Bap-5 (urea)	Bap-5 (dial)	AL
P.69	-3.399 1.441						
P.69 (urea)	-3.807 1.033	-2.828 2.012					
Bap-5 (PBS)	**-5.403** **-0.563**	-4.425 0.416	-4.016 0.824				
Bap-5 (urea)	**-5.306** **-0.466**	-4.327 0.513	-3.919 0.921	-2.323 2.517			
Bap-5 (dial)	**-5.309** **-0.469**	-4.330 0.510	-3.922 0.918	-2.325 2.515	-2.423 2.418		
Al	**-4.956** **-0.116**	-3.977 0.863	-3.569 1.271	-1.973 2.867	-2.070 2.770	-2.068 2.773	
GP	**-5.437** **-0.597**	-4.458 0.382	-4.050 0.790	-2.454 2.386	-2.551 2.289	-2.548 2.292	-2.901 1.939

There is an apparent discrepancy: R12 ACV was significantly more protective than P.69 and the same was apparent in R14. However when the data from R12, R13 and R14 were combined there was no significant difference between them (Table 10A). This is because of statistical fluctuations in the data being affected by the absence of P.69 in experiment R13.

The summary statistics for the two protective antigens and the AL control were produced. The median and confidence intervals are the most appropriate way to summarise the lung counts for the three preparations since the means are biased by the use of notional values (Table 10, part C).

Figure 41a: Combined results of total lung counts (log_{10}), with preparations ACV, P.69 and AL, from Experiments R12, R13 and R14. Each point represents one mouse. Counts of <100 or >$3x10^6$ were given notional values of 20 and 10^7 respectively and plotted as open symbols. The Minitab 'jitter' option was applied so as to spread the points horizontally to avoid coincidence.

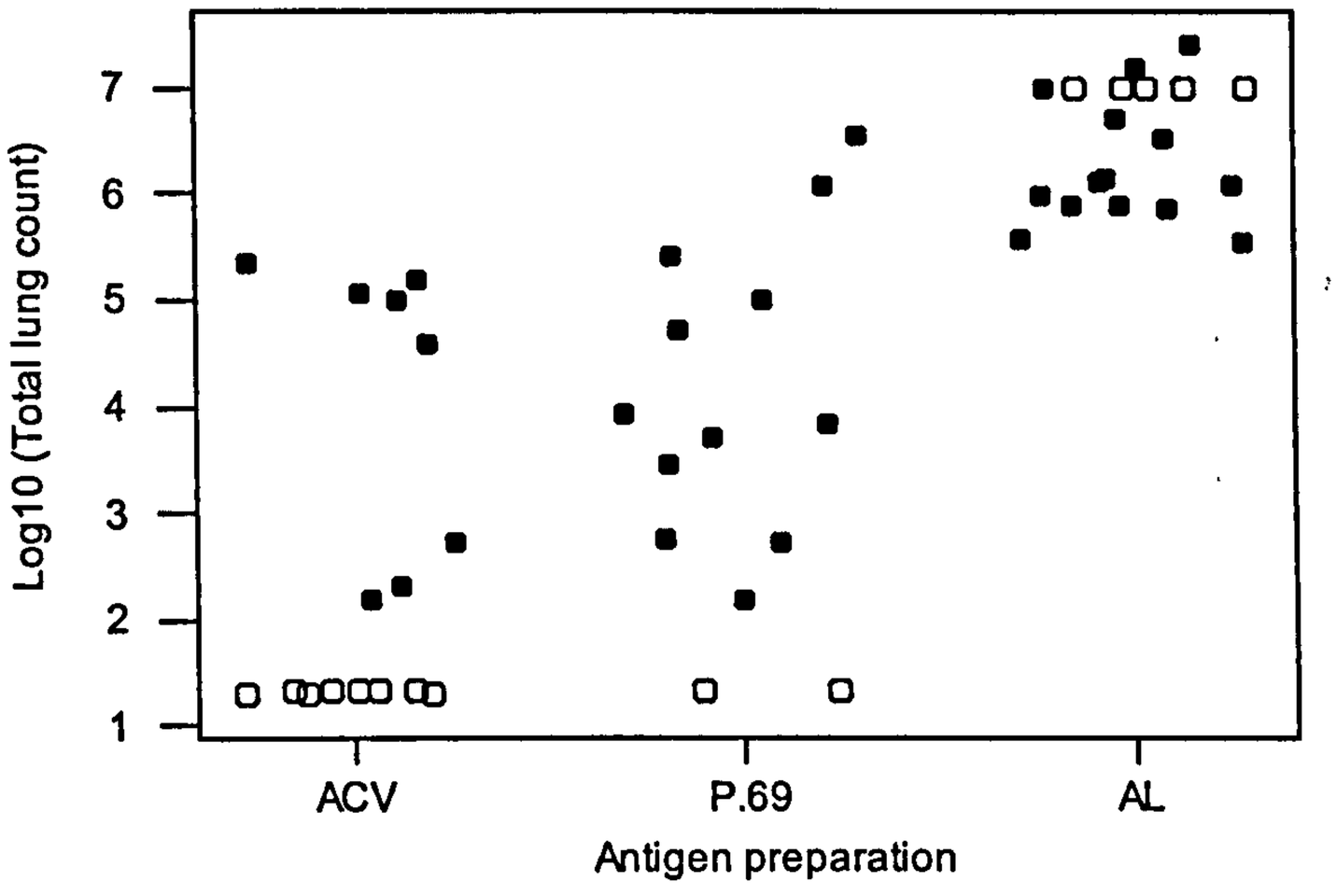


Figure 41b: Boxplots of the data in Fig 41a (above)

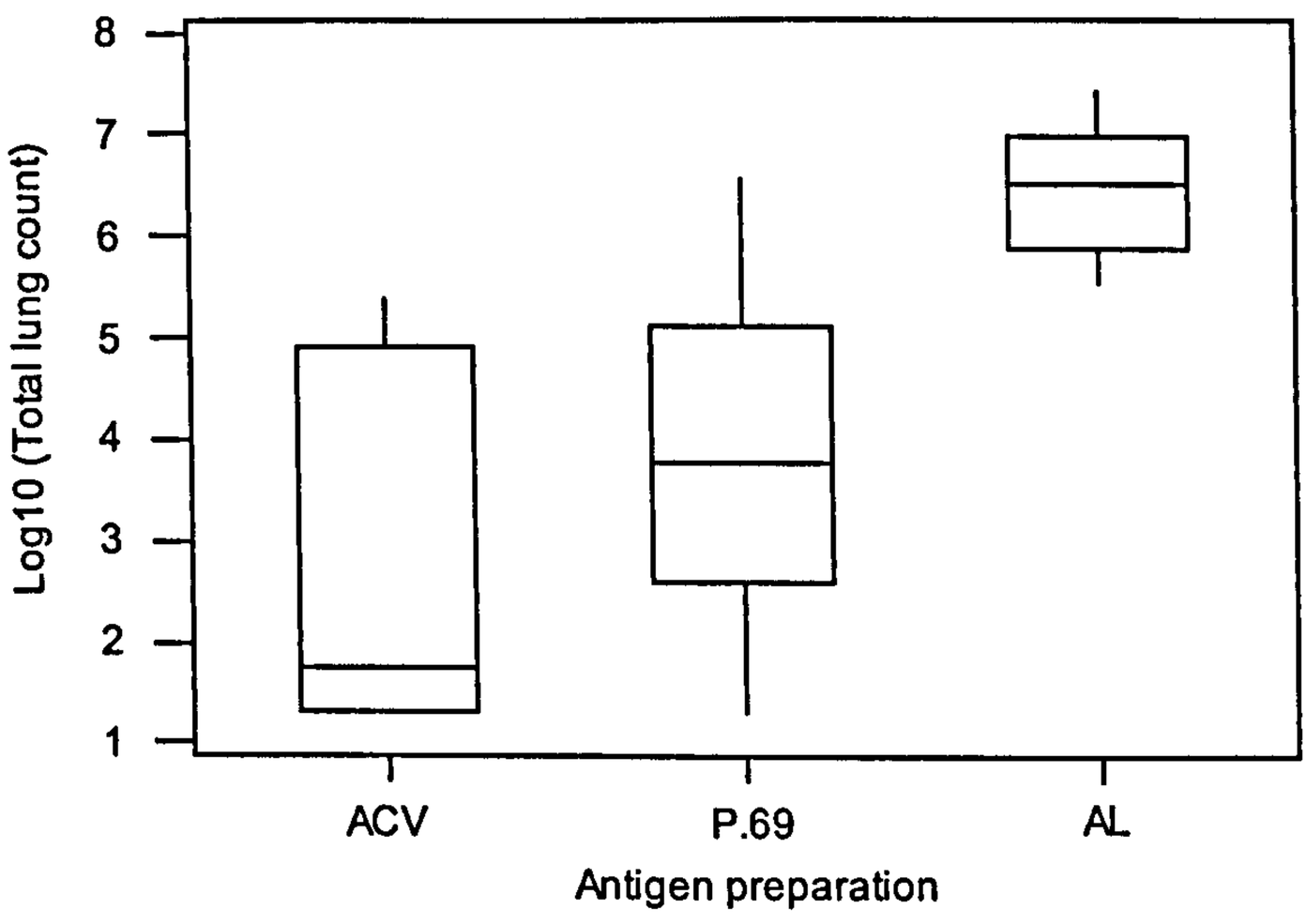


Table 10: Statistical analysis of the data in Fig 41.

Part A. Summarized output from Minitab of the *one-way analysis of variance* to determine whether, with each antigen preparation, there were significant differences between experiments.

Antigen preparation	P-value	Interpretation
ACV	0.238	Not significant
P.69	0.164	Not significant
AL	0.547	Not significant

These results justified the combination of the data for each antigen preparation, across the experiments, to give the point plots and boxplots in Fig. 41.

Part B. Analysis of variance and Tukey test on the three antigen preparations

B1. Primary output of the one-way Anova: Antigen preparation: P = 0.001 (Highly significant)

B2. Basic statistics and graphical summary (Mean and 95% Confidence intervals)

Antigen preparation	No. obs.	Mean	StDev	Mean $\log_{10}$ count and 95% CI
ACV	16	2.677	1.713	(----*---)
P.69	14	3.775	1.648	(----*----)
AL	19	6.471	0.600	(---*---)
Pooled StDev =		1.366		------+---------+---------+---------
Scale of Mean ($\log_{10}$ count)				➞3.0 4.5 6.0

B3. Tukey's pairwise comparisons of lung-count means ($\log_{10}$ values)
The table shows the 95% Confidence intervals for {*Column antigen mean* minus *Row antigen mean*}. Significant differences in mean counts are shown in **bold.**

	ACV	P.69
P.69	-2.310 0.114	
AL	**-4.918** **-2.670**	**-3.863** **-1.530**

Part C. Summary statistics: Mean, median, SEM and confidence intervals (CI) of *Total log_{10} lung counts* for each antigen preparation, combining the results of all the experiments

Antigen preparation	No. obs.	Mean	SEM	Median	95% CI of median
ACV	16	2.677	0.428	1.739	1.301, 4.696
P.69	14	3.775	0.441	3.763	2.672, 5.023
AL	19	6.471	0.138	6.531	5.985, 7.000

4.0 Discussion

4.1 Overview

A new member of the *B. pertussis* autotransporter family has been identified. A mis-primed PCR led to preliminary data which suggested the presence of a pertacin-like sequence in the *B. pertussis* genome. This open reading frame was identified using a combination of library screening and SSP-PCR techniques. Further characterisation of this open reading frame (*bap-5*) implicated the encoded protein (named Bap-5) as a virulence-regulated outer-membrane protein which appears to be processed in a similar manner to pertactin. Motif analysis suggests that Bap-5 may function as an adhesin. Construction of a *bap-5* mutant was undertaken but the final stages were not completed.

Expression constructs were created for Bap-5 and also the C-terminal domains of Prn, BrkA and Tcf which allowed large amounts of protein to be expressed and, in most cases, purified. The expression strains created will provide a source of protein which can be used in future mouse protection trials. Fusion of the pertactin C-terminus and linker region to a signal sequence allowed this domain to be targeted to the outer membrane. This construct may be of future use in display of heterologous proteins on the bacterial surfaces.

Preliminary experiments to examine the protective capacity of the autotransporter domains (C-terminal domains of Prn, Tcf, BrkA and Bap-5) in addition to Bap-5 and a specific portion of Bap-5 showed no protection in a murine model under the conditions tested.

4.2 Autotransporter characterisation

During the course of this study, a number of reports appeared in the scientific literature indicating that autotransporter proteins are present in a wide range of Gram-negative genera. This family of proteins has been established by comparison of the amino acid

sequences of their C-terminal domains. These domains are thought to allow self-promoted, extracellular secretion by forming a monomeric β–barrel. The most conserved amino acids between the autotransporters are those that traverse the outer membrane, as proposed by hydrophobicity and amphipathicity plots (Loveless and Saier 1997). The computer-based predictions for β–sheet formation are reputed to be accurate, although prediction of the number of sheets formed by a protein is significantly less so (Jose *et al.* 1995). The function of the mature N-terminal domains is varied and ranges from adhesion to proteolysis (Table 2). Recently, the autotransporter protein family has been reviewed, a signature sequence has been assigned, and the evolution and probable structure of these proteins evaluated (Loveless and Saier 1997; Henderson *et al.* 1998; Holland 1998). The signature sequence suggested by Loveless and Saier (1997), based on 18 known autotransporter proteins was:

[SATQRGH]-[AKLTQNF]-[FGATL]-[GPDSNKI]-[KRHNDQS]-X_3-[DPAQH]-
[QNDWKYFIT]-[ASTQN]-[LIVFGAS]-[NQHEGAIF]-[LIVA]-[NGSK]-[LIVMFY]-
[RKQSTG]-[LIFVYW]-[STAKRNV]-[FW]

This signature sequence is proposed to form the final beta sheet in the outer membrane and complete the predicted beta barrel. This sequence has also been implicated in outer membrane targeting.

The emergence of whole genome sequencing and subsequent analysis (genomics) has allowed confirmation of the autotransporter sequences and has allowed identification of further, related open reading frames in *E. coli* and, more recently, *B. pertussis* and *B. bronchiseptica* (section 1.7). The autotransporter secretion pathway has been named the type IV secretion pathway (Finlay and Falkow 1997).

4.3 Characterisation of *bap-5*

4.3.1 Sequence analysis of *bap-5*

4.3.1.1 Characterisation of the MR30 amplicon and identification of cosmid 3

As described in section 3.1, the amplification of the 3' end of *bap-5* (MR30) occurred due to mispriming by the oligonucleotides that were designed to anneal to part of the *prn* 3' sequence. The sequence of MR30 obtained was similar to, but different from, the predicted *prn* sequence (Figure 5).

From my investigation of the genome sequence, it is now apparent that the *B. pertussis* genome contains at least 10 genes which may be classified as autotransporter proteins on the basis of predicted amino acid homology (section 3.1.8). It is perhaps not surprising that the MR30 amplicon arose from a genomic template containing so many similar genes, especially as the primers used were directed towards the highly conserved putative processing site and the outer membrane localisation signal (Figure 3). The *bap-5* sequence (Figure 12) revealed that the MR30 primer binding sequences were very similar to those of pertactin.

To obtain the corresponding gene sequence within unamplified DNA, to avoid potential errors caused by DNA polymerase activity and therefore ensure accuracy of the sequence, a cosmid library was screened using part of the MR30 amplicon as a probe (P640) (section 3.1.2). One cosmid, cosmid 3, contained sequence identical to the MR30 amplicon. In addition, analysis of the flanking sequence revealed that MR30 did represent part of a larger putative ORF. Screening of a cosmid library using an amplicon as a Southern blot probe is a commonly used technique to aid the identification of genes. This is often a laborious task and several alternative approaches are currently available. In the case of *B. pertussis*, the genome sequence alleviates this need. In instances where

the genome sequence is not known, other approaches are now more favoured. Perhaps the most convenient is the use of an adapter tag to allow the construction of a molecular library. From such libraries, target genes can be amplified by PCR using only one gene-specific oligonucleotide in conjunction with an adapter-targeted primer. If a project similar to the one described was approached at the present time, such an approach would prove more convenient.

4.3.1.2 Obtaining the 3' sequence of *bap-5*

Unfortunately, the cosmid insert did not contain the 3' end of the putative ORF or encode a suitable translational stop codon, due to the *Eco*RI site (position 2341, Figure 5) being restricted during cosmid library construction. To obtain the 3' of the *bap-5* sequence and thus complete the putative open reading frame, attempts were made to re-screen the cosmid library using another Southern blot probe, P200, which was,like P640, derived from MR30 (Figure 4). This probe proved unsuitable and resulted in high background, probably due to non-specific hybridisations. This was possibly due to the high GC content of this region (65%).

Conventional Single Specific Primer-PCR (SSP-PCR) was also attempted using PEB1FOR16 with M13FOR or M13REV (Table 5). Although, on several occasions, the pUC ligated *B. pertussis* DNA resulted in plausible amplicons, the resultant sequences were always due to mispriming of the specific primer.

GenomeWalker™ SSP-PCR proved to be an effective way to establish the sequence at the 3' end of the putative open reading frame. Such approaches have major advantages over conventional techniques. This method obviates the requirement for a library stored at -80°C in *E. coli*, as 5 libraries (in microfuge tubes) are stored at -20°C. This technique allows rapid identification of sequences and is limited only by PCR elongation, which is in turn dependent on the distance between the restriction sites (adapter location) and the target site of the gene-specific primer. Libraries can be re-used approximately 80

times for identification of different flanking sequences. Resultant amplicons can be conveniently cloned using AT cloning techniques, which are often easier than cosmid subcloning.

The nucleotide sequence of *bap-5*, including some flanking sequence, and the predicted open reading frame was submitted to GenBank on the 11/08/1998 as a "putative autotransporter protein of *B. pertussis*" (accession number AF081494).

4.3.1.3 Detection of *bap-5* in the genomes of bordetellae

In addition to Southern blot identification of *bap*-5 in the wild-type *B. pertussis* genome of strains Taberman and BBC29 (Table 3), it was also possible to identify related sequences, that hybridised with the P640 probe in the *Eco*RI digested genomes of *B. bronchiseptica* and also *B. parapertussis* (Figure 16). Such bands were present at a different size in these species. The Prn and BrkA mutants (strains BBC30, a derivative of BBC29 wild type, and BP2041, respectively) of *B. pertussis* also showed hybridisations to restriction fragments of the same size as the wild type strains. The genes that are mutated in these strains represent the sequences with most similarity to the probe and it was considered that cross reaction to these genes might have been possible. The results suggested that the target DNA was not within *prn* or *brkA* and represented a distinct sequence. The presence of *bap-5* in the genomes of *B. pertussis* and *B. bronchiseptica* has since been confirmed by the genome sequencing projects (see Figure 42 for *B. bronchiseptica bap-5* homologue). The sequence in *B. pertussis* strain Tohama reported by the Sanger Center has 100% nucleotide identity to the *bap-5* nucleotide sequence obtained for this project using strain Taberman (BLAST search performed at http://www.sanger.ac.uk/Projects/B_pertussis). The fact that the *bap-5* sequence is identical in two *B. pertussis* isolates suggests functional importance. Also, a gene with high homology to Bap-5 has been identified in *B. bronchiseptica* strain RB50, which also infers commmon functional importance. The absence of a reaction of P640 to the *B. avium* genome was not surprising, as many other virulence factor genes,

including pertactin, are absent from this species and the envelope profile of *B. avium* is distinctly different to those of other bordetellae (Leyh and Griffith 1992).

Bordetellae harbour redundant genes, such as the flagellin genes in *B. pertussis and B. parapertussis* and the *brkA* gene in some *B. bronchiseptica* strains (Arico and Rappuoli 1987; Leigh *et al.* 1993; Rambow *et al.* 1998). The *brk* locus has been identified in all *Bordetella* species except *B. avium* and is expressed by all *B. pertussis* tested but only by some *B. bronchiseptica* strains, although its role in *B. bronchiseptica* is unclear (Rambow *et al.* 1998). Another example of silent genes is the pertussis toxin (ptx) operon which, although present in *B. bronchiseptica* and *B. parapertussis*, is not expressed. This is due to altered, defective, promoter sequences which prevent transcription (Arico and Rappuoli 1987). It was therefore important to confirm the expression of the open reading frame of Bap-5 at the transcriptional level (by RT-PCR, Figure 17) and also at the translational level (by Western blotting, Figure 18).

From the data presented in this thesis Bap-5 appears to be a protein of predicted molecular weight 79.5 KDa, and transcription is Bvg-regulated. The protein appears to be localised in the outer membrane and processed to fragments of apparent molecular weight, 30 KDa, and 65 KDa on SDS-PAGE of *B. pertussis* outer membrane preparations. The 30 KDa band is consistent with that predicted for the C-terminal portion of Bap-5 and it is feasible that the 65 KDa band represents a surface associated N-terminal portion (predicted size 49 KDa). However, size estimation of the N-terminus from SDS-PAGE is probably inaccurate. It is known that the related protein, P69, migrates anomolously on SDS-PAGE.

4.4 Bap-5 as an autotransporter

The similarities between the *B. pertussis* autotransporters obtained by performing a BLASTp search with the Bap-5 C-terminus amino acid sequence are summarised in Figure 14. The N-terminus of Bap-5 exhibits some sequence similarity to other

described *B. pertussis* autotransporters but not Tcf (Figure 15). The high sequence similarity between the Bap-5 C-terminus and other known autotransporters (Figure 15) strongly suggests that, on the basis of such homology, Bap-5 is another member of the autotransporter family. This high similarity is not surprising as these C-termini probably function to form a ß-barrel pore and allow translocation of the N-terminus to the cell surface. This similarity suggests that these proteins are closely related, and probably diverged following gene duplication, but with conservation of the C-terminal encoding regions. It is interesting to note the lack of similarity between the Tcf N-terminus and the other *B. pertussis* autotransporters (Figure 15). The N-terminal similarity between Prn, BrkA, Vag8 and Bap-5 is approximately 40% whereas the Tcf N-terminus possesses no significant similarity. This may imply a different origin for the 5' of the gene encoding Tcf and have implications for the function of Tcf compared to the other *B. pertussis* autotransporter proteins.

The role of the more highly conserved 20 amino acids at the C-termini of autotransporters must be established. It has been suggested for IgA protease and AIDA-I that they consitute a signal of outer membrane localisation and form the final ß-sheet in the outer membrane (Klauser *et al.* 1993; Suhr *et al.* 1996). Further evidence is provided by the Hap protein of *H. influenzae.* If the C-terminal three amino acids of the outer membrane localisation signal are mutated or removed, the protein is no longer detected in the outer membrane (Hendrixson *et al.* 1997).

4.4.1 The putative functional domains and motif analysis of Bap-5

The N-terminus of Bap-5 exhibits some homology to Prn and BrkA (Figure 15). These proteins all act as adhesins and BrkA also confers serum resistance to *B. pertussis* (Fernandez and Weiss 1994; Finn and Stevens 1995; Everest *et al.* 1996). Several motifs were identified in the predicted amino acid sequence of Bap-5. A potential integrin binding motif, arg-gly-asp (RGD), was identified within the N-terminus of Bap-5

(Figure 12). Such motifs are present in the N-terminal domains of Prn, BrkA and Tcf. (Figure 3). In Prn, the role of the RGD in adhesion and invasion has been studied using synthetic RGD-containing sequences in competitive binding experiments with HeLa cells. The results suggested that RGD in pertactin may play a role in adhesion and invasion (Leininger *et al.* 1992). Seemingly contradictory results were obtained when an RGD→RGE mutant of recombinant pertactin expressed by *E. coli* had no significant effect on adhesion to CHO or HepC cells (Everest *et al.* 1996). The position of the RGD within Bap-5 could be determined from structural analysis. Such experiments would be useful in determining a role for the RGD motif in Bap-5.

The presence of a glycosaminoglycan-binding motif, serine-glycine-serine-glycine (SGSG, position 1346, Figure 12), provides further evidence that Bap-5 may function as an adhesin and promote binding to glycosaminoglycans such as heparin. Such motifs (SGXG) are also present in BrkA and Tcf (Figure 3) although the role of these motifs in these proteins has still to be established experimentally. In addition, glycosaminoglycan-binding motifs have also been implicated in mediating resistance to complement, possibly by inhibiting polymerisation of the complement component C9 in a manner similar to that suggested for the Rck protein in *Salmonella typhimurium* (Hefferman *et al.* 1992; Hefferman *et al.* 1992; Fernandez and Weiss 1994). The serum protein vitronectin also has a glycosaminoglycan binding site which also inhibits C9 polymerisation (Millis *et al.* 1993). Again, the role of this motif in adhesion, and possibly complement resistance could be evaluated using synthetic peptides in competitive binding experiment and also by characterisation of a Bap-5 mutant lacking this motif.

A N-terminal leader (signal) sequence was not identified in Bap-5 using matrix prediction methods which are based on characterised signal sequences (Perlman and Halvorson 1983; Von Henge 1986). Prn, Tcf and Vag8 have typical signal sequences *ie.* those predicted to target the protein through the inner membrane by the *Bordetella* equivalent of the Sec pathway (Charles *et al.* 1989; Finn and Stevens 1995). It is likely,

from *E. coli* studies of autotransporters that the cleavage of a leader sequence, if present, occurs during translocation via the Sec-dependent pathway, or equivalent. This presumably accounts for the targetting of Prn, Tcf and Vag8 to and across the cytoplasmic membrane. In the case of BrkA, and possibly Bap-5, an alternative pathway must be utilised. For BrkA, BrkB appears to fulfil this role. It is possible that BrkB, or a functionally-related protein may be responsible for the translocation of Bap-5 across the cytoplasmic membrane. The use of a *B. pertussis* BrkB mutant to study the processing and export of Bap-5 (expression and cellular localisation studies) may provide evidence for the secretion pathway used by Bap-5. Alternatively, the *B. pertussis* genome sequence may aid identification of factors which may be responsible for the export of Bap-5. However, BLAST analysis of the genome database revealed no BrkB homologues (results not shown).

Several autotransporters, such as SepA and AIDA-I, appear to have atypical, longer signal sequences which include N-terminal extension sequences (Benjelloun-Toulmi *et al.* 1995). It has been suggested that such large atypical signal sequences present in autotransporters could be acting as an unfoldase in the periplasm (Holland 1998). Again, no such signal sequence could be identified within Bap-5. It is likely, therefore, that the export of Bap-5 across the cytoplasmic membrane is not mediated by a type II pathway.

4.4.2 Regulatory sequences adjacent to the proposed bap-5 open reading frame

The transcriptional start of *bap-5* was not determined during this study. Perhaps this could be achieved using 5'RACE (rapid amplification of 5' ends). This may aid the identification of the regulatory sequences upstream from the *bap-5* ORF. No typical Shine-Dalgarno sequence (Shine and Dalgarno 1975) was identified in the sequence 5' to the proposed open reading frame. Also, no BvgA dimer binding site was identified, which have been described upstream from the promoters of *prn, fha, bvg, ptx* and *cya* and takes the form TTTC(C/T)TA (Marques and Carbonetti, 1997; Kinnear *et al.*

1999). The BvgA dimer is thought to interact with RNA polymerase and regulates transcription.

Downstream from the *bap-5* proposed open reading frame there is a sequence which may represent a *rho*-independent terminator (Figure 13). There is also a putatitive *rho*-independent terminator for the gene encoding P.93 (Charles *et al.* 1989). In *P93*, this secondary structure is further from the stop codon and smaller than in *bap-5*. This potential hairpin structure at the 3' end of the mRNA could be responsible for causing the RNA polymerase to pause and therefore preventing further mRNA synthesis (Stryer 1988).

4.5 Identification of novel *B. pertussis* autotransporter proteins

The *B. pertussis* genome sequence (http://www.sanger.ac.uk/Projects/B_pertussis) was considered almost complete at the time of this analysis (finishing/gap closure taking place). However, according to the database, this sequence may contain errors and vector sequences and is not contigous. The amino acid sequences of the characterised *B. pertussis* C-terminal domains were BLAST searched against the genome data to identify potentially novel open reading frames for autotransporter (C-terminal) domains. The resultant coding sequences, and flanking sequences were then analysed using Genejockey II. The first methionine following an in-frame stop codon was used where appropriate. Figure 43 shows an alignment of the characterised autotransporters and the novel autotransporters Bap-6 - Bap-9 and Phg. Figure 44 shows a matrix comparison to highlight gross sequence similarity in the N and C-terminal domains. Phg has not yet been characterised, but has been deposited in GenBank (accession number AJ009835) as a cold shock protein of *B. pertussis*.

Several similarities can be identified from this preliminary information:

Bap-6: No start codon was identified within the novel N - terminus which contains forty four RR repeats. The significance of these repeats is not known. Bap-6 also contains an RGD motif.

Bap-7: No start codon was identified within the N-terminus. The N-terminus does have some homology to P.69 and contains an SGXG motif.

Bap-8: No ATG start codon was identified, although it is possible that a GTG start is present. The N-terminus has some homology with Tcf and P.69 and contains two PQP repeats.

Bap-9: No ATG start codon was identified. The N-terminus has some homology with Tcf and AIDA-I and contains an RGD and a SGXG motif.

The reason for *B. pertussis* possessing so many adhesins with similar proposed mechanisms of binding, such as those mediated by RGD and SGXG, is not clear. It has been suggested that it may be to allow "immune exclusion" by allowing the cells to avoid the effect of antibodies to specific adhesins which would prevent adhesion (Weiss, 1997).

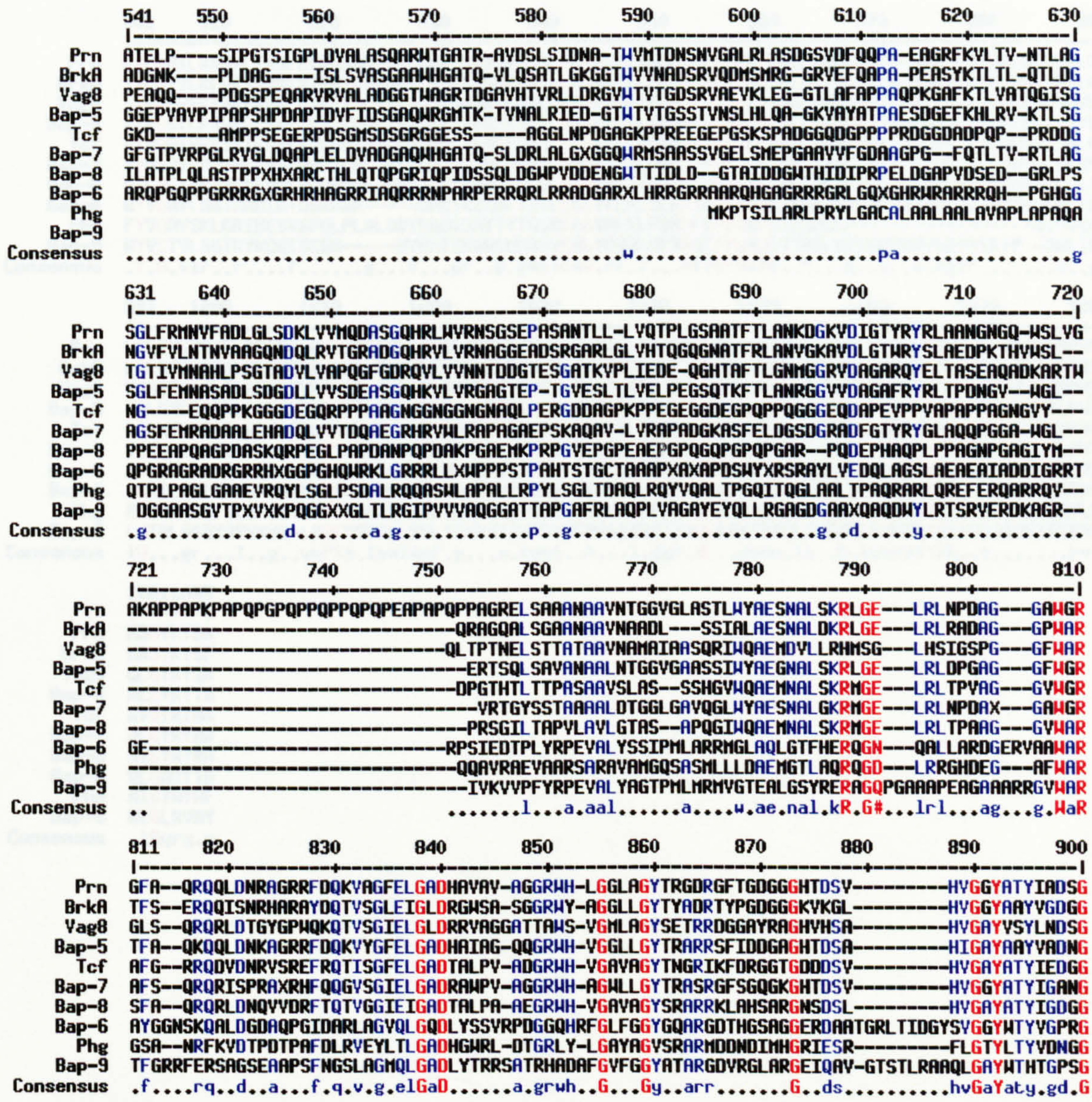
Figure 42: An alignment of the amino acid sequences of Bap-5 from *B. pertussis* (top) and *B. bronchiseptica* (bottom). Dots represent amino acid identity, lines represent amino acids with similar properties.

```
          10        20        30        40        50        60        70        80        90
          |         |         |         |         |         |         |         |         |
MCDTCRDDDGTSPSIRVQGGVVQGGMGANNVAVVATGSGKVAIENAELLGASGMYATFGAQVDMKGGRILAHNTNILGSQGYADGPYGGVVVTEDGQ
MCDTCRDDDGTSPSIRVRGGVVQGGVGANNVAVVATGPGKVAIENAELLGASGMYATFGAQVDMKGGRILAHNTNVIGSQDYADGPYGGVVLTEDGQ

  100       110       120       130       140       150       160       170       180       190
  |         |         |         |         |         |         |         |         |         |
VNLEGAKVSATGLGAAGLWLLGDKDTSPRASLRNTDVHGEVAAIALGFNGEANISGGSLSVEDGAVLTTLTPDAVEYYYDYALSMEHLPADAPLTPV
VNLEDVKVSATGLGAAGLWLLGDKDTSPRANLRNTDVHGEVAAIALGFNGEANISGGSLSVEDGAVLTTLTPDAVEYYYDYALSMEHPAADAPLTPV

      200       210       220       230       240       250       260       270       280       290
      |         |         |         |         |         |         |         |         |         |
RVTLSDGARASGETLIAHGGLLPMTLRLSSGVDARGDIVTLPPSAPPDSAEQPDAEPEPDAELEPDAAAQSDAKANARVMAQVDGGEPVAVPIPAPS
RVTLSDGARASGETLIAHGGLLPMTLRLSSGADARGDIVTLPPSAPPDSAEQPDAEPRT------DAAAQSDA•AIARVMAQVDGGEPXAXYDRALR

          300       310       320       330       340       350
          |         |         |         |         |         |
HPDAPIDVFIDSGAQWRGMTKTVNALRIEDGTWTVTGSSTVNSLHLQAGKVAYATPAESDGE----------------------------------
ISMPRSTCSSTVAPHGWASXTPSMRCASRIAPXTVTGSSQCTACSCRLAXVRTQTLAISERPFKHLPNQDPVRNLACTK•TPATDLNQCQLTGNVRP

                            360       370       380       390       400       410       420
                            |         |         |         |         |         |         |
--------------------FKHLRVKTLSGSGLFEMNASADLSDGDLLVVSDEASGQHKVLVRGAGTEPTGVESLTLVELPEGSQTKFTLANRGGV
SQPATRRYLYPGNRNGNLGKFKHLRVKTLSGSGLFEMNASADLSDGDLLVVSDEASGQHKVLVRGAGTEPTGVESLTLVELPKGSQAGFTLANRGGV
```

```
430        440        450        460        470        480        490        500        510        520
|          |          |          |          |          |          |          |          |          |
VDAGAFRYRLTPDNGVWGLERTSQLSAVANAALNTGGVGAASSIWYAEGNALSKRLGELRLDPGAGGFWGRTFAQKQQLDNKAGRRFDQKVYGFELG
•••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••
VDAGAFRYRLTPDNGVWGLERTSQLSAVANAALNTGGVGAASSIWYAEGNALSKRLGELRLDPGAGGFWGRTFAQKQQLDNKAGRRFDQKVYGFELG

  530        540        550        560        570        580        590        600        610        620
  |          |          |          |          |          |          |          |          |          |
ADHAIAGQQGRWHVGGLLGYTRARRSFIDDGAGHTDSAHIGAYAAYVADNGFYFDSTLRASRFENDFTVTATDAVSVRGKYRANGVGATLEAGKRFT
•••••••••••••••••••••••••••|•••••••••••••••••••••••••••••••••••••••••••••••|•••••••••••••••••••••
ADHAIAGQQGRWHVGGLLGYTRARRSFVDDGAGHTDSAHIGAYAAYVADNGFYFDSTLRASRFENDFTVTATDAVSIRGKYRANGVGATLEAGKRFT

      630        640        650        660        670        680        690        700        710     720
      |          |          |          |          |          |          |          |          |       |
LHDGWFVEPQSEVSLFHASGGTYRAANNLSVKDEGGTSAVLRLGLAAGRRIDLGKDRVIQPYATLSWLQEFKGVTTVRTNGYGLRTDLSGGRAELAL
••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••••
LHDGWFVEPQSEVSLFHASGGTYRAANNLSVKDEGGTSAVLRLGLAAGRRIELGKDRVIQPYATLSWLQEFKGVTTVRTNGYGLRTDLGGGRAELAL

          730        740        750
          |          |          |
GLAAALGRGHQLYTSYEYAKGNKLTLPWTFHLGYRYTW
••••••••••||••••••••••••••••••••••••|•
GLAAALGRGHKFYTSYEYAKGNKLTLPWTFHLGYRYAW
```

Figure 43: An alignment of the characterised autotransporters with those predicted from the *B. pertussis* genome sequence during the course of this study. Red amino acids represent identity and blue repesent similar amino acids.

```
                1       10        20        30        40        50        60        70        80        90
                |-------+---------+---------+---------+---------+---------+---------+---------+---------|
            Prn
           BrkA MYLDRFRQCPSSLQIPRSAWRLHALAAALALAGMARLAPAAAQAPQPPVAGAPHAQDAGQEGEFDHRDNTLIAVFDDGVGINLDDDPDEL
           Vag8                                                                                  MAGQAR
          Bap-5
            Tcf
          Bap-7
          Bap-8
          Bap-6
            Phg
          Bap-9
      Consensus                                                                                      ......

                91      100       110       120       130       140       150       160       170       180
                |-------+---------+---------+---------+---------+---------+---------+---------+---------|
            Prn         MNMSLSRIVKAAPLRRTTLAMA-LGALGAA---PAAHADWNNQSIVKTGERQHGIHIQGSDPGGVRTASGTTIKVSGRQ--AQ
           BrkA GETAPPTLKDIHISVEHKNPMSKPAIGVR-VSGAGRALTLAGSTIDATEGGIPAVVRRGGTLELDGVTVAGGEGMEPMTVSDAGSRLSVR
           Vag8 GWYGAGGRHPIHFQISAGAALMLGLLDVAGAAAVTAAQRIDGGAAFLGDVAIATTKASEHGINVTGRTAE--VRVTGGTIRTSGNQ--AQ
          Bap-5
            Tcf
          Bap-7
          Bap-8
          Bap-6
            Phg
          Bap-9
      Consensus ............................. ..................................................... ..

                181     190       200       210       220       230       240       250       260       270
                |-------+---------+---------+---------+---------+---------+---------+---------+---------|
            Prn GILLENP-AAELQFRNGSVTSSGQLSDDGIRRFLGTVTVKAGKL--VADHATLANVGDTWDD----------------DGIALYVAGEQA
           BrkA GGVLGGE-APGVGLVRAAQGGQASIIDATLQSILGPALIADGGSISVAGGSIDMDMGPGFPPPPPPLPGAPLAAHPPLDRVAAVHAGQDG
           Vag8 GLRVGTENAPDNTALGASVFLQNLIIETSGTGALGVSVHEPQGGGGTRLSMSGTTVRTRGDDSFALQLSGPASA--TLNDVALETAGQQA
          Bap-5                                                                                  MCDTCRDDDGTSP
            Tcf
          Bap-7
          Bap-8
          Bap-6                                                                                            PQ
            Phg
          Bap-9
      Consensus ....... ..................................................................... .............

                271     280       290       300       310       320       330       340       350       360
                |-------+---------+---------+---------+---------+---------+---------+---------+---------|
            Prn QASIADSTL-----QGAGGVQIERGANVTVQRSAI-VDGGLHIGALQSLQPED-LPPSRVVLRDTNVTAVPASGAPAAVSVLGASELTLD
           BrkA KVTLREVALRAHGPQATGVYAYMPGSEITLQGGTVSVQGDDGAGVVAGAGLLDALPPGGTVRLDGTTVSTDGANTDAVLVRGDAARAEVV
           Vag8 PAVVLWQGAQLNAQGLVVQVNGAGVSAIHAQDAGSFTLSGSDITARGLEVAGIYVQEGMQGTLTGTRVTTQGDTAPALQVEDAGTHVSMN
          Bap-5 SIRVQGGVVQGGMGANNVAVVATGSGKVAIENAELLGASGMYATFGAQVDMKGGRILAHNTNILGSQGYADGPYGGVVVTEDG--QVNLE
            Tcf                                                                                MHIYGNMNRATPCRGAVRALALA
          Bap-7
          Bap-8
          Bap-6 PYRQRGHDPDGRRRRRLARRHGGFRRQQPVDQPRIDHDIGQRRARHLGAQRRQYRDAGRQQRQHPHPGRQRARHRHQRPGQPRPEQRRHH
            Phg
          Bap-9
      Consensus ..........................................................................................

                361     370       380       390       400       410       420       430       440       450
                |-------+---------+---------+---------+---------+---------+---------+---------+---------|
            Prn GGHITGGR--AAGVAAMQGAVVHLQRATIRRGDALAGG-AVPGGAVPGGAVPGGFGPGGFGPVLDG--WYGVDVSGSSVELAQSIVEAPE
           BrkA NTVLRTAKSLAAGVSAQHGGRVTLRQTRIETAGAGAEGISVLGFEPQSGSGPASVDMQGGSITTTGNRAAGIALTHGSARLEGVAVRAEG
           Vag8 GGALSTSGANSPAAWLLAGGSAQFRDTVLRTVGEASHGVDV---------AAHSEVELAHAQVRADGQGAHGLVVTRSSAMVRAGSL-YES
          Bap-5 GAKVSATGLGAAGLWLLGDKDTSPRASLRNTDVHGEVAAIALGFNGEANISGGSLSVEDGAVLTTLTPDAVEYYYDYALSMEHLPADAPL
            Tcf LLGAGMWTLSPPSAWALKLPSLLTDDELELVLPTGMSLEDFKRSLQESAPSALATPPSSSPPVAKPGPGSVAEAPSGSGHKDNPSPPVVG
          Bap-7
          Bap-8
          Bap-6 RQRHRRQGRVSARRQPRGQRTGQRRHHPCARRQQQRHIRPRRRACEHHQRQRLSCPRREPARRAHPQRSLVCAARAERQRYLHQRRLPAR
            Phg
          Bap-9
      Consensus ..........................................................................................

                451     460       470       480       490       500       510       520       530       540
                |-------+---------+---------+---------+---------+---------+---------+---------+---------|
            Prn LGA-AIRVGRGARVTVPGGSLSAPHGWVIETGGARRFAPQA---APLSITLQ-AGAHAQGK--ALLYRVLPEPV-KLTLTGGADAQGDIV
           BrkA SGSSAAQLANGTLVVSAGSLASAQSGAISVTDTPLKLMPGALASSTVSVRLT-DGATAQGGNGVFLQQHSTIPV-AVALESGALARGDIV
           Vag8 TGDGAAALLESGHLTVDGSVVHGHGAAGLEVDGES----------NVSLL-NGARLSSDQPTAIRLIDPRSVLNLDIKDRAQLLGDIA
          Bap-5 T---PVRVTLSDGARASGETLIAHGGLLPMTLRLSSGVDARGDIVTLPPSAPPDSAEQPDAEPEPDAELEPDAAAQSDAKANARVMAQVD
            Tcf VGPGMAESSGGHNPGVGGGTHENGLPGIGKVGGSAPGPGGLGRNDENSESSLNPGTLGPSPGPDTSTGSGPDAGMASGAGSTSPGASGGA
          Bap-7                                                 RWAGRPEWRVAGGAQLLRQTTPVPVRLVLRGAAVAQGDVVRAPETAPEKD
          Bap-8                                                                                VGLRSHLPELPTPSRCHE
          Bap-6 ARRGRPRHGRLHGPPGHGHADPAHRLGHRRPGRRRRGGQPCVSGGQRHGGQPVRQLPHPDHARRRLALDLGRRVHRKRGPAHRHILSCRH
            Phg
          Bap-9
      Consensus ................g.....................................................................
```


541 550 560 570 580 590 600 610 620 630
Prn ATELP----SIPGTSIGPLDVALASQARWTGATR-AVDSLSIDNA-TWYMTDNSNVGALRLASDGSVDFQQPA-EAGRFKVLTV-NTLAG
BrkA ADGNK----PLDAG-----ISLSVASGAAWHGATQ-VLQSATLGKGGTWVVNADSRVQDMSMRG-GRVEFQAPA-PEASYKTLTL-QTLDG
Vag8 PEAQQ----PDGSPEQARVRVALADGGTWAGRTDGAVHTVRLLDRGVWTVTGDSRVAEVKLEG-GTLAFAPPAQPKGAFKTLVATQGISG
Bap-5 GGEPVAVPIPAPSHPDAPIDVFIDSGAQWRGMTK-TVNALRIED-GTWTVTGSSTVNSLHLQA-GKVAYATPAESDGEFKHLRV-KTLSG
Tcf GKD------AMPPSEGERPDSGMSDSGRGGESS-----AGGLNPDGAGKPPREEGEPGSKSPADGGQDGPPPPRDGGDADPQP--PRDDG
Bap-7 GFGTPVRPGLRVGLDQAPLELDVADGAQWHGATQ-SLDRLALGXGGQWRMSAASSVGELSMEPGAAVVFGDAAGPG--FQTLTV-RTLAG
Bap-8 ILATPLQLASTPPXHXARCTHLQTQPGRIQPIDSSQLDGWPVDDENGWTTIDLD--GTAIDDGWTHIDIPRPELDGAPVDSED--GRLPS
Bap-6 ARQPGQPPGRRRGXGRHRHAGRRIAQRRRNPARPERRQRLRRADGARXLHRRGRRAARQHGAGRRRGRLGQXGHRWRARRRQH--PGHGG
Phg MKPTSILARLPRYLGACALAALAALAVAPLAPAQA
Bap-9
Consensus .....w.....pa.....g
631 640 650 660 670 680 690 700 710 720
Prn SGLFRMNVFADLGLSDKLVVMQDASGQHRLWVRNSGSEPASANTLL-LVQTPLGSAATFTLANKDGKVDIGTYRYRLAANGNGQ-WSLVG
BrkA NGVFVLNTNVAAGQNDQLRVTGRADGQHRVLVRNAGGEADSRGARLGLVHTQGQGNATFRLANVGKAVDLGTWRYSLAEDPKTHVWSL--
Vag8 TGTIVMNAHLPSGTADYLVAPQGFGDRQVLVVNNTDDGTESGATKVPLIEDE-QGHTAFTLGNMGGRVDAGARQYELTASEAQADKARTW
Bap-5 SGLFEMNASADLSDGDLLVVSDEASGQHKVLVRGAGTEP-TGVESLTLVELPEGSQTKFTLANRGGVVDAGAFRYRLTPDNGV--WGL--
Tcf NG----EQQPPKGGGDEGQRPPPAAGNGGNGGNGNAQLPERGDDAGPKPPEGEGGDEGPQPPQGGGEQDAPEVPPVAPAPPAGNGVY---
Bap-7 AGSFEMRADAALEHADQLVVTDQAEGRHRVWLRAPAGAEPSKAQAV-LVRAPADGKASFELDGSDGRADFGTYRYGLAQQPGGA-WGL--
Bap-8 PPEEAPQAGPDASKQRPEGLPAPDANPQPDAKPGAEMKPRPGVEPGPEAEPGPQGQPGPQPGAR--PQDEPHAQPLPPAGNPGAGIYM--
Bap-6 QPGRAGRADRGRRHXGGPGHQWRKLGRRRLLXWPPPSTPAHTSTGCTAAAPXAXAPDSWYXRSRAYLVEDQLAGSLAEAEAIADDIGRRT
Phg QTPLPAGLGAAEVRQYLSGLPSDALRQQASWLAPALLRPYLSGLTDAQLRQYVQALTPGQITQGLAALTPAQRARLQREFERQARRQV--
Bap-9 DGGAASGVTPXVXKPQGGXXGLTLRGIPVVVAQGGATTAPGAFRLAQPLVAGAYEYQLLRGAGDGAAXQAQDWYLRTSRVERDKAGR--
Consensus .g.....d.....a.g.....p..g.....g..d.....y.....
721 730 740 750 760 770 780 790 800 810
Prn AKAPPAPKPAPQPGPQPPQPPQPQPEAPAPQPPAGRELSAAANAAVNTGGVGLASTLWYAESNALSKRLGE---LRLNPDAG---GAWGR
BrkA ----------QRAGQALSGAANAAVNAADL---SSIALAESNALDKRLGE---LRLRADAG---GPWAR
Vag8 ----------QLTPTNELSTTATAAVNAMAIAASQRIWQAEMDVLLRHMSG---LHSIGSPG---GFWAR
Bap-5 ----------ERTSQLSAVANAALNTGGVGAASSIWYAEGNALSKRLGE---LRLDPGAG---GFWGR
Tcf ----------DPGTHTLTTPASAAVSLAS--SSHGVWQAEMNALSKRMGE---LRLTPVAG---GVWGR
Bap-7 ----------VRTGYSSTAAAALDTGGLGAVQGLWYAESNALGKRMGE---LRLNPDAX---GAWGR
Bap-8 ----------PRSGILTAPVLAVLGTAS--APQGIWQAEMNALSKRMGE---LRLTPAAG---GVWAR
Bap-6 GE----------RPSIEDTPLYRPEVALYSSIPMLARRMGLAQLGTFHERQGN---QALLARDGERVAAWAR
Phg ----------QQAVRAEVAARSARAVAMGQSASMLLLDAEMGTLAQRQGD---LRRGHDEG---AFWAR
Bap-9 ----------IVKVVPFYRPEVALYAGTPMLMRMVGTEALGSYRERAGQPGAAAPEAGAAARRGVWAR
Consensus .. ..l...a.aal.....a....w.ae.nal.kR.G#...lrl...ag...g.WaR
811 820 830 840 850 860 870 880 890 900
Prn GFA--QRQQLDNRAGRRFDQKVAGFELGADHAVAV-AGGRWH-LGGLAGYTRGDRGFTGDGGGHTDSV---------HVGGYATYIADSG
BrkA TFS--ERQQISNRHARAYDQTVSGLEIGLDRGWSA-SGGRWY-AGGLLGYTYADRTYPGDGGGKVKGL---------HVGGYAAYVGDGG
Vag8 GLS--QRQRLDTGYGPWQKQTVSGIELGLDRRVAGGATTAWS-VGMLAGYSETRRDGGAYRAGHVHSA---------HVGAYVSYLNDSG
Bap-5 TFA--QKQQLDNKAGRRFDQKVYGFELGADHAIAG-QQGRWH-VGGLLGYTRARRSFIDDGAGHTDSA---------HIGAYAAYVADNG
Tcf AFG--RRQDVDNRVSREFRQTISGFELGADTALPV-ADGRWH-VGAVAGYTNGRIKFDRGGTGDDDSV---------HVGAYATYIEDGG
Bap-7 AFS--QRQRISPRAXRHFQQGVSGIELGADRAWPV-AGGRWH-AGWLLGYTRASRGFSGQGKGHTDSV---------HVGGYATYIGANG
Bap-8 SFA--QRQRLDNQVVDRFTQTVGGIEIGADTALPA-AEGRWH-VGAVAGYSRARRKLAHSARGNSDSL---------HVGAYATYIGDGG
Bap-6 AYGGNSKQALDGDAQPGIDARLAGVQLGQDLYSSVRPDGGQHRFGLFGGYGQARGDTHGSAGGERDAATGRLTIDGYSVGGYWTYVGPRG
Phg GSA--NRFKVDTPDTPAFDLRVEYLTLGADHGWRL-DTGRLY-LGAYAGVSRARMDDNDIMHGRIESR---------FLGTYLTYVDNGG
Bap-9 TFGRRFERSAGSEAAPSFNGSLAGMQLGADLYTRRSATRHADAFGVFGGYATARGDVRGLARGEIQAV-GTSTLRAAQLGAYWTHTGPSG
Consensus .f....rq..d..a...f.q.v.g.elGaD.....a.grwh..G...Gy..arr.....G..ds.....hvGaYaty.gd.G

```
                901       910       920       930       940       950       960       970       980       990
                |--------+---------+---------+---------+---------+---------+---------+---------+---------|
Prn             FYLDATLRASRLENDFKVAGSDG-YAVKGKYRTHGVGASLEAGRRFTHADGWFLEPQAELAVFRAGGGAYR-AANGLRVRDEGGSSVLGR
BrkA            YYLDTVLRLGRYDQQYNIAGTDG-GRVTADYRTSGAAWSLEGGRRFELPNDWFAEPQAEVMLWRTSGKRYR-ASNGLRVKVDANTATLGR
Vag8            SYVDGVVKYNRFRHGFDIRTTDL-KRVDAKHRSHGLGALLRGGRRIDIDGGWYVEPQASVAWFHAGGSRYE-ASNGLRVRADGAHSWVLR
Bap-5           FYFDSTLRASRFENDFTVTATDA-VSVRGKYRANGVGATLEAGKRFTLHDGWFVEPQSEVSLFHASGGTYR-AANNLSVKDEGGTSAVLR
Tcf             FYMDGIVRVSRIRHAFKVDDAKG-RRVRGQYRGNGVGASLELGKRFTWPGAWYVEPQLEVAAFHAQGADYT-ASNGLRIKDDGTNSMLGR
Bap-7           VYADATLRASRFENSFDAPGWAG-RTVSGSYRANGVGVTLEAGRRLALDRHWFVEPQAELAWFRAGGGTYT-ASNGLRIEDDGGTSLQAR
Bap-8           FYLDGIVRVNRYEHDFRADGQRG-ARVTGKYRANGIGLSLETGRRFTWAGDWFVEPQVEVALFRSGGADYT-ASNGVRVDVASTKSLLGR
Bap-6           WYVDAVLANTWMDIDTDSKAGR-----DADTRGQAFTASLESGYPLALSERWTLEPQAQLIYQHTRVDGFSDAVSEVRIRDD--NALTAR
Phg             FYVDAVSKLGRIDESVSFDLPLGLGDYDDDISHTTYTGSAEAGYHFKLPQRWFVEPQAQVIYSRSS----------------QTSVQGR
Bap-9           WYVDTVLAGTRYKQQTSSSA-----HVGATSRGWGMMASVEAGYPWQLNPRWQIEPQAQLVYQQLGIANGADRVSSVSYKTP--DALTGR
Consensus       .Y.D.vlr..r....f......g...v...yr..g.gaslEaG.rf.l...WfvEPQa#v..f...g..y..a.nglr.......s..gR

                991   1000      1010      1020      1030      1040      1050      1060      1070      1080
                |--------+---------+---------+---------+---------+---------+---------+---------+---------|
Prn             LGLEVGKRIELAGGRQVQPYIKASVLQEFDGAGTVHTNGI-AHRTELRGTRAELGLGMAAALGRGHSLYASYEYSK--G--PKLAMPWTF
BrkA            LGLRFGRRIALAGGNIVQPYARLGWTQEFKSTGDVRTNGI-GHAGAGRHGRVELGAGVDAALGKGHNLYASYEYAA--G--DRINIPWSF
Vag8            AGAEAGRQMRLANGNIVEPYARLGWAQELGADNAVYTNGI-RHVTRSRGGFAEARVGVGALLGKRHALYADYEYAK--G--ARFEAPWTL
Bap-5           LGLAAGRRIDLGKDRVIQPYATLSWLQEFKGVTTVRTNGY-GLRTDLSGGRAELALGLAAALGRGHQLYTSYEYAK--G--NKLTLPWTF
Tcf             LGLHVGRQFDLGDGRVVQPYMKLSWVQEFDGKGTVRTNDI-RHKVRLDGGRTELAVGVASQLGKHGSLFGSYEYAK--G--SRQTMPWTF
Bap-7           VGAQAGRRFDLRGGAVVQPYAQLSWVQELKGVSTVRTNGI-AHRTDLGAGRVELGLGVAAALGKGHNLYASYEYAH--G--PRLSLPWTY
Bap-8           AGLQVGRKLDLGNGKLVQPYAKLSWLQEFDGVGKVRTNDI-GHDVKLRGGRAELDLGVAAALGRHSSLFASYEYSK--G--SRLTIPWSF
Bap-6           LGARLQGEYAAA-AQVWRPYAALNFWRTFSGENTVVLGED-SIDTRRGATSLELAAGASVTLARSLALYGRLAYAT--SIDSQYLRGASA
Phg             AGVRAGRDFTLAGGATLRPYVSASYLHEFSHDDSVDFGGK-SYDAELPGSRWQLGAGAALDVG-AHRAYADLRYGH--G--ANISQDLSL
Bap-9           LGTRLAGQYAYRKAQ-LRPFMGVSLLHDLTGXXXXTXXGVHSVRASXQXTXVDLKAXVDTQLGKSVGLWXQVGYGKSVGSGDGSDXGWSA
Consensus       lG...gr...l..g..vqP%a.lswlqef.g...v.tngi..h...l.ggr.#L..gvaa.lg..h.lyasyeYak..g.......pws.

                10811088
                |------|
Prn             HAGYRYSW
BrkA            HAGYRYSF
Vag8            QLGYRYSW
Bap-5           HLGYRYTW
Tcf             HVGYRYAW
Bap-7           QLGYRYAW
Bap-8           HVGYRYAW
Bap-6           QLGMRYTW
Phg             NIGYAYRF
Bap-9           NLGLRVAY
Consensus       .lGyry.w
```

Figure 44: A matrix to demonstrate amino acid similarities between the *B. pertussis* autotransporter domains described to date and those that have been identified in the Sanger *B. pertussis* genome sequence. The C-terminal and N-terminal domains were chosen on the basis of known or putative cleavage sites. All similarities were determined by BLASTp analysis and are displayed as the % positive (identical and similar) amino acids / extent of the overlap. The extent of the overlap is the window of amino acids which have been analysed. NS represents no significant similarity.

Bap-5: GenBank Acc No. AF081494

Prn: GenBank Acc No. AJ006152

BrkA: GenBank Acc No. U12276

Tcf: GenBank Acc No. U16754

Vag8: GenBank Acc No. U90124

	Bap-5	**Prn**	**BrkA**	**Tcf**	**Vag8**
Bap-6	N: NS C: 40/280	N: NS C: 43/279	N: NS C: 39/256	N: NS C: 41/277	N: NS C: 38/280
Bap-7	N: 53/224 C: 69/278	N: 48/276 C: 74/278	N: NS C: 63/278	N: NS C: 63/278	N: 42/196 C: 57/266
Bap-8	N: NS C: 68/279	N: NS C: 57/279	N: NS C: 64/279	N: 43/138 C: 78/279	N: NS C: 57/279
Bap-9	N: NS C: 42/282	N: NS C: 41/291	N: NS C: 41/280	N: NS C: 42/287	N: NS C: 37/284
Phg	N: NS C: 43/280	N: NS C: 45/279	N: NS C: 43/279	N: NS C: 43/279	N: NS C: 39/268

4.6 Processing of autotransporter proteins

The mechanism of processing of the bordetellae autotransporters is unknown at present, although autoproteolytic release from the cell surface does not seem possible as no protease catalytic sites have been reported within the characterised autotransporter proteins. However, a eukaryotic baculovirus expression system was able to process P.93 to P.69 which implies a general protease or autoproteolytic activity (Charles *et al.* 1993). There is evidence that Tcf is not autoproteolytic, and requires a *B. pertussis* specific serine protease (Chen *et al.* 1998). Such evidence was obtained from studying the processing of Tcf in *Vibrio cholerae.* In Bap-5, no site with homology to the IgA protease serine protease active site motif (GDSGSPLF) (Brenner 1988; Bachovchin *et al.* 1990) was identified.

The gonococcal IgA protease is processed in *E. coli* by the envelope serine protease OmpT (Klauser *et al.* 1992). Therefore, the most likely candidate protease for processing other autotransporter proteins might be a serine protease in the the cell envelope. However, the processing of autotransporter proteins in *E. coli* does not necessarily reflect the native situation. For example, although OmpT is responsible for the processing in *E. coli* of IgA protease (from *N. gonorrhoeae*) this process is thought to be autoproteolytic in *N. gonorrhoeae.* AIDA-I (from *E. coli*) appears autoproteolytic, as processing occurs in *E.coli* strains deficient in ompT, ompP and DegP (Suhr *et al.* 1996). Other *E. coli* autotransporters, EspP and Pet, and the Hap protein of *Haemophilus influenzae* contain serine protease motifs which are thought to be processing targets (Brunder *et al.* 1997; Hendrixson *et al.* 1997; Eslava *et al.* 1998; Navarro-Garcia *et al.* 1999). Some autotransporter proteins, such as VacA (from *H. pylori*), may possess species-specific transport mechanisms as they are not secreted or processed in *E. coli* (Manetti *et al.* 1995).

In *E. coli,* OmpT is responsible for degrading IcsA (from *S. flexneri*) which prevents the intracellular spread of the bacteria (Nakato *et al.* 1993). However, in *S. flexneri*, a

candidate protease, *Shigella* outer membrane protease A (SopA) has been implicated in the polar localisation and processing of IcsA (Egile *et al.* 1997). This protease exhibits 60% amino acid homology with OmpT and OmpP proteases of *E. coli.* It has been suggested for IcsA that the C-terminal domains may have high affinity for components characteristic of domains at the pole of the bacterial cell surface (Steinhauer *et al.* 1999). The same may be true for bordetellae autotransporters and raises the possibility of a focal point for adhesins at the bacterial surface. Another *Shigella* autotransporter, SepA, contains a serine protease catalytic site although this site is not necessary for cleavage of the SepA precursor and no protease activity has been demonstrated for this protein (Benjelloun-Toulmi *et al.* 1995).

Several *E. coli* envelope proteases have been characterised and such information may aid identification of proteases responsible for autotransporter processing in other bacterial species (Lazdunski 1989; Strauch *et al.* 1989; Bayeyx and Georgiou 1991; Meerman and Georgiou 1994; Pallen and Wren 1997). To investigate candidate enzymes for such proteolysis, an *E. coli* expression system could be used. For example, *E. coli* strain BL21(DE3) (see Table 3) which is deficient in ompT was used as a host strain in which to express Prn. It would appear that this protease is not responsible for the processing of pertactin as both P.30 and P.69 was recovered following expression of pertactin at low levels in this *E. coli* strain (M. Roberts, personal communication). In addition, several other protease mutants, deficient in one or more envelope proteases were donated by G. Georgiou at the University of Texas at Austin and are described elsewhere (Bayeyx and Georgiou 1991; Meerman and Georgiou 1994). Such mutants were intended for the study of the processing mechanism of recombinant bordetellae autotransporter proteins and it was hoped that candidates in *B. pertussis* could be identified from such information. Due to time constraints, and the emergence of the genome sequence, these experiments were not performed in the present study.

In light of the of the *B. pertussis* genome sequence, perhaps a more rational approach would include identification and characterisation of potential proteases that are present

in the *B. pertussis* cell envelope. The advent of bioinformatics will allow rapid analyisis of the genome sequence data (Patterson and Handel 1998). The PEDANT internet site (http://pedant.mips.biochem.mpg.de/) provides information regarding categorised theoretical open reading frames from the *B. pertussis* genome sequence data which may suggest candidate proteases. Such proteases could be mutated (knocked-out) and the processing mechanisms of autotransporter proteins evaluated by cell fractionation and Western blotting, using specific antisera to the N-termini and the C-termini, of strains expressing autotransporter proteins. The cellular location of the N-terminus, C-terminus and unprocessed protein domains in such strains could be identified by Western blotting with the specific antibodies that were raised to recombinant domains during this study.

4.6.1 Outer membrane processing of Bap-5

The presence of the consensus cleavage site in Bap-5, as in the other autotransporters, suggests the potential for cleavage and potential release from the cell surface. It is perhaps difficult to understand why the organism would produce adhesins that are released from the cell surface. The regulation of such release may provide information regarding the pathogenesis of pertussis. The control of processing may represent an important mechanism for alteration of adhesive properties of *B. pertussis* at different stages of infection *i.e.* to enable adhesion to different sites in the host, to allow transmission to a new host, or cell invasion and subsequent intracellular survival may be dependent on such release. Future studies on mutants which have an altered processing site and are unable to correctly process the precursor protein may provide vital information regarding autotransporter protein processing.

4.7 Expression of recombinant domains

4.7.1 Expression construct design

To express Bap-5, a specific portion of Bap-5 (NTS) and the C-terminal portions of Bap-5, Prn, BrkA and Tcf, primers were designed to allow directional cloning into either pET11a, pET33b or pQE-60. With the exception of the BrkA C-terminus, all expression constructs contained a His_6 tag (N-terminal for pET33b and C-terminal for PQE-60 constructs) for ease of purification. Where necessary, translational start or stop codons were introduced (Table 6). Prior to cloning into the expression vectors, the amplicons were first AT cloned and sequenced to check the correct target sequence and the fidelity of the amplification.

The amplification of the C-terminal encoding DNA was difficult, probably due to the high GC content of these domains (72.1% for Prn, 70.55 for BrkA and 68.76 for Tcf). This is higher than the overall GC content of the *B. pertussis* which is reported to be 66 %. Initial attempts to amplify these domains with proof reading DNA polymerase probably failed for this reason. Early in this study, DMSO was considered to be an aid to difficult PCR, but since then, many manufacturers have made available tools which facilitate amplification of such difficult templates. For example, reagents such as Q solution (Qiagen) are now available, which outperform DMSO for amplification of GC rich templates. All C-terminal encoding regions could be amplified under the same conditions.

4.7.2 Expression of the autotransporter domains

All of the expression constructs that are described in this study could be induced to produce large amounts of recombinant protein. It was also possible to confirm the C-terminus and NTS proteins by Western blotting with mAbs to either the His_6 tag or

with specific mAbs (sections 3.2.4 and 3.2.5). The over-expression of the autotransporter domains, and also Bap-5, appeared to be independent of the IPTG concentration within the range tested (0.2 - 1.5 mM), and in all cases a significant quantity of protein was present within the host cells at 1 h post induction. Also, there appeared to be no expression of recombinant protein prior to induction, which was evident both by Coomassie-blue staining and by more sensitive Western blotting techniques.

All of the recombinant proteins expressed during this study appeared to form inclusion bodies within the cell. Such structures enabled very high level expression of potentially-damaging hydrophobic membrane proteins within the cells. For structural and functional analysis, other cell fractions could be analysed and purification under native conditions may be desirable. The inclusion bodies were solubilised in 8 M urea (Buffer B, appendix I), which provided suitable material for affinity purification under denatured conditions, where appropriate. The solubilised inclusion bodies proved to be stable at -20°C for at least 18 months. However, when stored at 4°C for more than one week, degradation of the pertactin C-terminus was apparent on SDS-PAGE (other domains not tested, results not shown).

4.7.3 Affinity purification of His_6 proteins

Further purification of the C-terminal domains and NTS was possible. On most occasions it was possible to produce very pure protein on either a large or small scale. The Ni-NTA microspin columns provided a very rapid and convenient technique to purify recombinant protein from the urea-solubilised inclusion bodies. The FPLC was more time consuming, although a much greater yield was possible. A preliminary, unsuccessful attempt was made to purify the whole Bap-5, although optimisation of this was not attempted as urea extracts provided protein of sufficient purity for the mouse protection tests. No Western blot was performed to confirm the presence of the His_6 tag on Bap-5 and it is possible that the protein produced was not as intended.

However, the DNA sequence of the construct was confirmed as correct and the Bap-5 protein was identified at approximately its predicted MW.

Further mouse protection tests and structural studies could be performed now that the recombinant proteins can be readily expressed and purified in large amounts. At present, structural analysis, including X-ray crystallography, of the purified C-terminal domains, is currently being undertaken in collaboration with the Protein Crystallography Unit at the University of Glasgow. The pertactin C-terminus would complete the structure of unprocessed pertactin (P.93) and would be the first autotransporter structure to be resolved. Such information would also test the hypothesis that the C-terminal domain forms a beta-barrel in the outer membrane and conclusions from such evidence could be applied to other autotransporters (Maurer *et al.* 1999).

4.7.4 Heterologous antigen display

The pertactin C-terminus/linker region was cloned into an arabinose-inducible vector which encodes a signal sequence, pBAD/gIII. Expression was tightly controlled by the concentration of arabinose in the medium and resulted in a very high yield of pertactin C-terminus/linker, BADPCTa (43KDa) in the *E. coli* outer membrane fraction. Such control may be a useful factor in the expression of recombinant proteins that are toxic to the host cell. An additional protein in the outer membrane was identified which may be the processed form of the pertactin C-terminus, BADPCTb (29 KDa) which remains in the outer membrane following cleavage within the linker .

This is the first evidence that the C-terminus of a *B. pertussis* autotransporter domain alone is capable of associating with the outer membrane (when expressed in conjuction with a signal sequence). To a much lesser extent, a construct of AIDA-I fused to a leader (signal) sequence was capable of reaching the outer membrane (Suhr *et al.* 1996). The lack of heat modifiability of BADPCTa when solubilised at room temperature and 100°C and run on SDS-PAGE may reflect incorrect folding in the outer membrane and

its subsequent interaction with SDS (Figure 35). The heat modifiability of the smaller, BADPCTb protein may reflect correct conformation in the membrane without the constraints of the linker region. The nature of the interaction of the BADPCT proteins at the membrane needs to be examined further. Use of immunogold-labelled monoclonal antibodies against specific portions of this protein, perhaps BBO5 (section, 3.3.3.2), in conjunction with electron microscopy, would help to clarify the topology of the C-terminus in the membrane. Alternatively, a whole cell ELISA, with antibodies such as BBO5, may detect surface exposure of the linker (and therefore future passenger domains).

Several attempts have been made to characterise the minimum autotransporter domain *ie.* the minimum number of residues that are essential for correct targetting and translocation of the passenger domains (Miyazaki *et al.* 1989; Klauser *et al.* 1993; Suhr *et al.* 1996; Maurer *et al.* 1999). Such attempts have highlighted the importance of the linker region, which appears to have 2 functions. The C-terminus of the linker region is proposed to traverse the pore formed and inhibit leakage, whereas the N-terminus of the linker anchors the passenger to the cell surface until cleavage from the cell occurs.

The system described in this study (section 3.3) is based on the C-terminus (pore forming domain) and linker region of pertactin. The linker region was approximated and designed to include the $(PQP)_5$ region. The expression and processing of pertactin in both *E. coli* and *Salmonella spp.* has been characterised (Fairweather *et al*, 1990; Lipscombe *et al.* 1991; Strugnell *et al.* 1992; Charles *et al.* 1994; Everest *et al.* 1996). Therefore a potential exists for using the pertactin C-terminus is in the development of future *Salmonella* based vaccines, including those used for mucosal immunisation. Maurer *et al.* (1997) expressed a heterologous passenger protein, either CtxB or a five amino acid peptide antigen tag, which was fused to the AIDA-I C-terminus. Using this system, expression of the protein as approximately 5% of total cell protein was achieved from a constitutive promoter (Maurer *et al.* 1997). The vector system used here (pBAD/gIII) allowed excellent control of expression level and gave a yield of

approximately 5% of whole-cell protein (Figure 32b), although no heterologous passenger domain was included and this may affect the eventual level of expression achieved.

4.7.4.1 Apparent molecular weights of BADPCTa and BADPCTb

On SDS-PAGE (with or without boiling during sample preparation), the BADPCTa protein migrated at apparently 43 KDa , which is 15 % higher than the predicted size of 37.1 KDa. The presence of the linker region in the BADPCTa protein appears to cause the abberant migration of this protein and to abolish the heat modifiability seen with the smaller BADPCTb.

The migration of BADPCTb in SDS-PAGE is dependent on the temperature used to prepare the sample prior to gel loading. The effect of solubilisation temperature on the SDS-PAGE migration of certain "heat-modifiable" outer membrane proteins has been described elsewhere (Inouye and Yee 1973; Schnaitman 1973; Armstrong and Parker 1986; See and Jackowski 1990). When the outer-membrane preparation was prepared at room temperature, the BADPCTb protein was present as a 27 KDa protein. When the sample was boiled this protein migrated more slowly, at approximately 30 KDa. This observation is consistent with the findings of Parton and Wardlaw, who found that two major outer membrane protein in the outer membranes of *B. pertussis* (now confirmed as the Tcf C-terminus and BrkA C-terminus) migrated more slowly on SDS-PAGE following boiling of the outer-membrane preparations (Parton and Wardlaw 1975). More recently, Maurer *et al.* (1999) found that AIDA-I C-terminal domain migrated more slowly following boiling although the authors' discussion appears to contradict their findings. The issue of heat modifiability has been discussed previously and it is apparent that different proteins are modifiable in different ways. Generally, the electrophoretic mobility of proteins on SDS-PAGE is a balance between the shape and size of the protein and the amount of SDS bound (Maurer *et al.* 1999). Maurer *et al.* (1999) also proposed that the increase in the apparent molecular weight of the AIDA-I

C-terminal domain following boiling is consistent with the presence of ß-sheets in this protein which are only completely denatured by boiling in the presence of SDS and the same modifiability is seen with *E. coli* major porins. The results presented in Figure 35 are consistent with these findings. It is difficult to understand why the BADPCTa protein, which is presumed to be BADPCTb plus the linker region does not appear to be heat modifiable. It is possible that the presence of the linker region prevents the normal folding of the C-terminal domain which is thought to confer heat modifiability. The IgA protease carboxy terminal domain is protease resistant which further supports the notion that this region is embedded within the outer membrane and therefore not accessible (Klauser *et al.* 1993). The predicted and apparent weight of the P30 domain (purified as described in section 3.2.3.2), run in parallel (Figure 35) is 29.8 KDa.

4.7.4.2 Future development and potential constraints

The next step in the development of this vector may be to introduce a multiple cloning site between the linker and the signal sequence which would allow convenient cloning of passengers domains. Alternatively, the *Nco*I and *Xho*I vector-encoded sites (Figure 31) may serve as suitable sites for the directional cloning of passenger genes. In addition, the signal sequence and BADPCT encoding sequences could be amplified together from this vector and introduced into an alternative, more suitable vector for use in *Salmonella* mucosal vaccines. Alternatively, the whole cells expressing the BADPCT proteins could be investigated as vaccines against *B. pertussis* as the immunodominant $(PQP)_5$ repeat regions are encoded by the linker and may therefore afford protection.

The inclusion of the linker region appears to have enabled processing in *E. coli* of the 43 KDa protein (BADPCTa) to a putative processed form in the outer membrane of 29 KDa, (BADPCTb). If the processing of P.93 in *E. coli* could be better defined, it could potentially be controlled. There are two approaches to enable such control - the use of altered linker regions (see below) or, alternatively, the use of *E. coli* host strains deficient in specific envelope proteases. The latter option may be better than alteration of

passenger domains, which may have altered properties (Klauser *et al.* 1992; Anderson *et al.* 1997).

Alteration of the linker region may include incorporation of a synthetic protease site which could be engineered into the vector for this purpose and which could allow very controlled release of the passenger domains(s) from the cell surface. Alternatively the protease responsible for P.93 processing could be co-expressed in a regulated manner. Such a fusion protein which utilises the HylA transport system has been evaluated and the release of passenger protein was increased by co-expressing OmpT (Hanke *et al.* 1992). Such control may simplify purification of recombinant proteins from the cell surface or be of use in vaccine formulation. Perhaps a vaccine vector which can express a variety of antigens on the cell surface and allow the controlled release of some factors may be useful in the development of vaccines that can protect against a variety of diseases simultaneously.

An application of a *B. pertussis* adhesin has recently been demonstrated. The immunogenicity of a liposome-delivered antigen was enhanced when FHA was included in the liposome (Poulain-Godefroy *et al.* 1998). This effect was probably due to FHA adhesion to the mucosal tissues. The expression of heterologous adhesins on cell surfaces in similar systems could also be investigated and used for either drug or antigen targetting to specific tissues.

4.8 Protective properties of autotransporters

No protection was evident from the any of the recombinant proteins - whole Bap-5, NTS (Bap-5 specific portion) and the C-terminal domains of Prn, BrkA, Tcf and Bap-5 under the stated conditions. The mouse protection tests were validated by the confirmation that a Smithkline Beecham acellular vaccine and, to a lesser extent, purified native P.69 were protective, and the *E. coli* cell lysate and alhydrogel alone controls were not protective according to the mouse lung count data. The mouse weight gains did

not prove to be a useful indicator of infection by the challenge strains or of protection by the vaccine, due to the age of the mice and the time of the challenge.

The protective properties of mature pertactin (P.69) have been well documented (section 1.10.5). P.69 is an important component of the new range of multicomponent acellular vaccines and, therefore, the N-termini of other autotransporter proteins such as mature Bap-5, with significant sequence similarity and possibly similar virulence functions, may also be useful immunising agents under appropriate conditions.

The urea extracts containing recombinant *B. pertussis* autotransporter proteins, were chosen for the mouse protection tests as they provided a readily available source of relatively pure protein. It is known, and is confirmed in this study, that *E. coli* proteins are not protective against *B. pertussis* challenge in a mouse model. A major advantage of testing recombinant proteins is that no other *B. pertussis* candidate antigens were present in the vaccine doses, unlike earlier native vaccine preparations, many of which would have been contaminated by trace amounts of pertussis toxin, which is known to be protective and, in non-protective doses, confers protective activity on other *B. pertussis* components in the mouse intracerebral challenge test.

The abundance and homogeneity of the *B. pertussis* autotransporter proteins, in particular the C-terminal domains, suggests that these proteins may warrant further consideration as components in acellular vaccine preparations. Although the recombinant proteins tested within this preliminary study suggested a lack of protection against infection with *B. pertussis*, it may be possible to obtain protection under different circumstances and many other variables remain to be explored. For example, inclusion of the autotransporter proteins in outer membrane vesicles, or combined with other virulence factors, especially with a trace amount of pertussis toxin. The importance of presenting *B. pertussis* antigens in the correct form has been studied and it appears that the best response occurs in mouse models when outer membrane complexes or microspheres are used (Hamstra *et al.* 1995; Shahin *et al.* 1995). Hamstra

et al. (1995) reported that an outer membrane complex (protein-detergent micelles) containing a 32 KDa protein (now confirmed as Tcf C-terminus) was protective in an intracerebral mouse protection tests. The 92 KDa protein described by Hamstra *et al.* (1995) which has now been confirmed as Vag8 was protective only when non-protetctive levels of pertussis toxin were added.

Also, the experimental design, such as the vaccine and challenge doses and routes of administration of the vaccine, different adjuvant formulations or different experimental models such as the coughing rat model could be investigated (Hall *et al.* 1994). Monji *et al.* (1986) suggested a role for a 30 KDa protein in *B. pertussis* outer membrane preparations which potentiated an immune response to *Haemophilus* type B capsular polysaccharide, which suggests a role of the C-terminal domains as adjuvants.

4.9 The potential for rational drug design

The abundance, highly conserved nature and importance of the autotransporter beta barrel in *B. pertussis* and other Gram-negative species may provide the opportunity to prevent the translocation of many different virulence factors to the cell exterior. The potential may exist for the development of antibacterial agents which specifically disrupt the autotransporter export mechanism. Structural information, especially regarding the autotransporter beta domains would help to rationalise such drug design. The autotransporters of many important pathogens could potentially be targetted and the diseases prevented.

5.0 References

Ad Hoc Group, **(1988)** Placebo controlled trial of two acellular vaccines in Sweden. *Lancet* **1**, 955-966

Anderson, C. L., Matthey-Dupraz, A., Missiakas, D., & Raina, S. **(1997)**. A new *Escherichia coli* gene, *dbsg*, encodes a periplasmic protein involved in disulphide bond formation, required for recycling dsba/dsbb and dsbc proteins. *Molecular Microbiology*, **26**(1), 121-132.

Anderson, R., Dougan, G., & Roberts, M. **(1996)**. Delivery of the pertactin/P.69 polypeptide of *Bordetella pertussis* using an attenuated *Salmonella typhimurium* vaccine strain: expression levels and immune response. *Vaccine*, **14**(14), 1384-1390.

Anwar, H. **(1991)**. Surface-exposed antibody-accessible outer membrane proteins of *Bordetella pertussis*. *Canadian Journal of Microbiology*, **37**, 590-593.

Aoyama, T., Takeuchi, Y., Gotto, A., Iwai, H., Murase, Y., & Iwata, T. **(1992)**. Pertussis in adults. *American Journal of Diseases in Children*, **146**, 163-166.

Arico, B., & Rappuoli, R. **(1987)**. *B. Parapertussis* and *B. Bronchiseptica* contain transcriptionally silent pertussis toxin genes. *Journal of Bacteriology*, **169**(6), 2847-2853.

Armstrong, S. K., & Parker, C. D. **(1986)**. Heat-modifiable envelope proteins of *Bordetella pertussis*. *Infection and Immunity*, **54**(1), 109-117.

Bachovchin, W. W., Plaut, A. G., Flentke, G. R., Lynch, M., & Kettner, C. A. **(1990)**. Inhibition of igai proteases from *Neisseria gonorrhoeae* and *Haemophilus influenzae* by peptide prolyl boronic bonds. *Journal of Biological Chemistry*, **265**, 3738-3743.

Bassinet, L., Gueirard, P., Maitre, B., Housset, B., Gounon, P., & Guiso, N. **(2000)**. Role of adhesins and toxins in invasion of human tracheal epithelial cells by *Bordetella pertussis*. *Infection and Immunity*, **68**(4), 1934-1941.

Bayeyx, F., & Georgiou, G. **(1991)**. Construction and characterisation of *Escherichia coli* strains deficient in multiple secreted proteases: Protease III degrades high-molecular weight substrates *in vivo*. *Journal of Bacteriology*, **173**(8), 2696-2703.

Benjelloun-Toulmi, Z., Sansonetti, P. J., & Parsot, C. **(1995)**. Sepa, the major extracellular protein of *Shigella flexneri*: autonomous secretion and involvement in tissue invasion. *Molecular Microbiology*, **17**(1), 123-135.

Benz, I., & Schmidt, M. A. **(1992)**. AIDA-1, the adhesin involved in diffuse adherence of the diarrhoeagenic *Escherichia coli* strain 2787 (O126:H27), is synthesised via a precursor molecule. *Molecular Microbiology*, **6**(11), 1539-1546.

Biska, J. B., Galan, J. E., & Falkow, S. **(1993)**. Signal transduction in the mammalian cell during bacterial attachment and entry. *Cell*, **73**, 920.

Bordet, J., & Gengou, O. **(1906)**. Le microbe de la colueluche. *Ann Inst Pasteur*, **23**, 415-419.

Brennan, M. J., Li, Z. M., Cowell, J. L., Bisher, M. E., Steven, A. C., Novotny, P., & Manclark, C. R. **(1988)**. Identification of a 69-kilodalton nonfimbrial protein as an agglutinogen of *Bordetella pertussis*. *Infection and Immunity*, **56**(12), 3189-3195.

Brenner, S. **(1988)**. The molecular evolution of genes and proteins: a tale of two serines. *Nature*, **334**, 528-530.

Brown, D. R., & Parker, C. D. (**1987**). Cloning of the filamentous hemaglutinin of *Bordetella pertussis* and its expression in *Escherichia coli*. *Infection and Immunity*, **55**(1), 154-161.

Brownlie, R. M., Coote, J. G., & Parton, R. (**1986**). Complementation of mutations in *Escherichia coli* and *Bordetella pertussis* by *B. Pertussis* DNA cloned in a broad-host range cosmid vector. *Journal of General Microbiology*, **132**, 3221-3229.

Brunder, W., Schmidt, H., & Karch, H. (**1997**). Espp, a novel extracellular serine protease of enterohaemorrhagic *Escherichia coli* 0157:H7 cleaves human coagulation factor V. *Molecular Microbiology*, **24**(4), 767-778.

Burns, D. L. (**1999**). Biochemistry of type IV secretion. *Current Opinion in Microbiology*, **2**, 25-29.

Capiau, C., Carr, S. A., Hemling, M. E., Plainchamp, D., Conrath, K., Desmons, P., Permanne, P., & Petre, J. O. (**1990**). Purification, characterisation, and immunological evaluation of the 69-kDa outer membrane protein of *Bordetella pertussis*. In C. R. Manclark (Ed.), Proceedings of the Sixth International Symposium on Pertussis, (pp. 75-85). DHSS Publication Number (FDA) 90-164.

Charles, I., Fairweather, N., Pickard, D., Beesley, J., Anderson, R., Dougan, G., & Roberts, M. (**1994**). Expression of the *Bordetella pertussis* P.69 pertactin adhesin in *Escherichia coli*: fate of the carboxy-terminal domain. *Microbiology*, **140**, 3301-3308.

Charles, I. G., Dougan, G., Pickard, D., Chatfield, S., Smith, M., Novotny, P., Morrissey, P., & Fairweather, N. F. (**1989**). Molecular cloning and characterisation of protective outer membrane protein P.69 from *Bordetella pertussis*. *Proceedings of the National Academy of Science USA*, **86**, 3554-3558.

Charles, I. G., Li, L. J., Strugnell, R., Beesley, K., Romanos, M., Dougan, G., Novotny, P., Heron, I., Jensen, M., Manclark, C. R., Brennan, M. J., & Fairweather, N. F. **(1990)**. Repeat sequence motifs constitute the immunodominant regions of the P.69 protein, pertactin from *Bordetella pertussis* : comparison with repeat sequences from *Bordetella parapertussis* and *Bordetella bronchiseptica.* In C. R. Manclark (Ed.), Proceedings of the Sixth International Symposium on Pertussis, (pp. 136-140). DHSS Publication Number (FDA) 90-164.

Charles, I. G., Rogers, B., Musgrave, S., Peakman, T. C., Chubb, A., Fairweather, N. F., Dougan, G., & Roberts, M. **(1993)**. Expression of the P.69 pertactin from *B. Pertussis* in a baculovirus/insect cell expression system: protective properties of the recombinant protein. *Research in Microbiology*, **144**, 681-690.

Chen, I., Finn, T. M., Yanqing, L., Guoming, Q., Rappuoli, R., & Pizza, M. **(1998)**. A Recombinant Live Attenuated Strain of *Vibrio cholera* Induces Immunity against Tetanus Toxin and *Bordetella pertussis* Tracheal Colonistion Factor. *Infection and Immunity*, **66**(4), 1648-1653.

Cherry, J. D. **(1997)**. Coparative efficacy of accelular pertussis vaccines: an analysis of recent trials. *Journal of Pediatric Infectious Disease*, **1997**(16(Supplement)), 90-96.

Cherry, J. D., Brunnel, P. A., Golden, G. S., & Karzon, D. T. **(1988)**. Report on the task force on pertussis and pertussis immunization. *Pediatrics*, **81**, 939.

Cobb, B. D., & Clarkson, J. M. **(1994)**. A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods. *Nucleic Acids Research*, **22**(18), 3801-3805.

Cookson, B. T., Vandamme, P., Carlson, L. C., Larson, A. M., Sheffield, J. V. L., Kersters, K., & Spach, D. H. **(1994)**. Bacteraemia caused by a novel *Bordetella* species, "*B. Hinzii*". *Journal of Clinical Microbiology*, **32**(10), 2569-2571.

Coote, J. G. **(1991)**. Antigenic switching and pathogenicity: environmental effects on virulence gene expression in *Bordetella pertussis*. *Journal of General Microbiology*, **137**, 2493-2503.

Coote, J.G. **(1992)**. Structural and functional relationships among the RTX toxin determinants of Gram-negative bacteria. *FEMS Microbiology Reviews*, **88**, 137-162.

Cotter, P. A., & Miller, J. M. **(1996)**. Genetic analysis of the *Bordetella*-host interaction. *Annals of the New York Acadamy of Sciences*, **797**, 65-76.

Cotter, P. A., Yuk, M. H., Mattoo, S., Akerley, B. J., Boschwitz, J., Relman, D. A., & Miller, J. F. **(1998)**. Filamentous hemagglutinin of *Bordetella bronchiseptica* is required for efficient establishment of tracheal colonisation. *Infection and Immunity*, **66**(12), 5921-5929.

Cover, T. L. **(1996)**. The vacuolating cytotoxin of *Helicobacter pylori*. *Molecular Microbiology*, **20**(2), 241-246.

Cowan, S. W., Schrimer, T., Rummel, G., Steiert, M., Ghosh, R., Pauptit, R. A., Jansonius, J. N., & Rosenbusch, J. P. **(1992)**. Crystal structures explain functional properties of two *E. Coli* porins. *Nature*.

Craig, F. F., Lackie, J. M., Parton, R., & Freer, J. H. **(1988)**. Interaction of *Bordetella pertussis* virulence components with neutrophils: effect on chemiluminescence induced by a chemotactic peptide and by intact bacteria. *Journal of General Microbiology*, **134**, 2201-2211.

Cullinane, L. C., Alley, M. R., Marshall, R. B., & Manktelow, B. W. **(1987)**. *Bordetella parapertussis* from lambs. *New Zealand Veterinary Journal*, **35**, 175.

Cundell, D.R., Kanthakumar, K., Taylor, G. W., & Goldman, W. E. (**1994**). The effect of tracheal cytotoxin from *Bordetella pertussis* on human neutrophils function *in vitro*. *Infection and Immunity*, **62**, 639-643.

d'Enfert, C. (**1993**). Yet another chaperone? *Trends in Microbiology*, **1**(5), 161-162.

De Ley, J., Segers, P., Kersters, K., Mannheim, W., & Lievens, A. (**1986**). Intra- and intergenic similarities of the *Bordetella* ribosomal ribonucleic acid cistrons: Proposal for a new family, *Alcaligenaceae*. *International Journal of Systematic Bacteriology*, **36**(3), 405-414.

Dinh, T., Paulsen, I. T., & Saier, M. H. J. (**1994**). A family of extracytoplasmic proteins that allow transport of large molecules across the outer membranes of Gram-negative bacteria. *Journal of Bacteriology*, **176**(13), 3825-3831.

Domenighini, M., Relman, D., Capiau, C., Falkow, S., Prugnola, A., Scarlato, V., & Rappuoli, R. (**1990**). Genetic characterisation of *Bordetella pertussis* filamentous haemagglutinin: a protein processed from an unusually large precursor. *Molecular Microbiology*, **4**(5), 787-800.

Egile, C., d'Hauteville, H., Parsot, C., & Sansonetti, P. J. (**1997**). Sopa, the outer membrane protease responsible for polar localistion of icsa in *Shigella flexneri*. *Molecular Microbiology*, **23**(5), 1063-1073.

Eldering, G., & Kendrick, P. (**1938**). *Bacillus para-pertussis*: A species resembling both *Bacillus pertussis* and *Bacillus bronchisepticus* but identical with neither. *Journal of Bacteriology*, **35**, 561-572.

Emsley, P., Charles, I. G., Fairweather, N. F., & Isaacs, N. W. (**1996**). Structure of *Bordetella pertussis* virulence factor P.69 pertactin. *Nature*, **381**, 90-92.

Emsley, P., Mcdermott, G., Charles, I. G., Fairweather, N. F., & Isaacs, N. W. (**1994**). Chrystallographic characterisation of pertactin, a membrane-associated protein from *Bordetella pertussis*. *Journal of Molecular Biology*, **235**(2), 772-773.

Enfors, S. O. (**1992**). Control of *in vivo* proteolysis in the production of recombinant proteins. *TIBTECH*, **10**, 310-315.

Eslava, C., Navarro-Garcia, F., Czeczulin, J. R., Henderson, I. R., Cravioto, A., & Nataro, J. P. (**1998**). Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infection and Immunity*, **66**(7), 3155-3163.

Everest, P., Li, J., Douce, G., Charles, I., Deazavedo, J., Chatfield, S., Dougan, G., & Roberts, M. (**1996**). Role of the *Bordetella pertussis* P.69/pertactin protein and the P.69/pertactin motif in the adherence to and invasion of mammalian cells. *Microbiology*, **142**, 3261-3268.

Ewanowich, C. A., Leininger, E., Kenimer, J. G., & Peppler, M. S. (**1990**). Mechanisms of *Bordetella pertussis* invasion of hela 229 Cells. In C. R. Manclark (Ed.), Proceedings of the Sixth International Symposium on Pertussis, (pp. 106-114). DHSS Publication Number (FDA) 90-164.

Ezzell, J. W., Dobrogosk, W. J., Kloos, W. E., & Manclark, C. R. (**1981**). Phase-shift markers in *Bordetella*: alterations in envelope proteins. *The Journal of Infectious Diseases*, **143**(4), 562-569.

Fairweather, N. F., Makoff, A. J., Oxer, M. D., Ballantine, S., Roberts, M., Dougan, G., & Charles, I. G. (**1990**). P.69 Pertactin: high-level expression and purification from *Escherichia coli* and protective properties. In C. R. Manclark (Ed.), Proceedings of the

Sixth International Symposium on Pertussis, (pp. 275-280). DHSS Publication Number (FDA) 90-164.

Falkow, S. (1991). Bacterial entry into eukaryotic cells. *Cell*, 65, 1099-1102.

Femlee, T., Pellet, S., & Welch, R. A. (1985). *Escherichia coli* hemolysin is released extracellularly without cleabage of a signal peptide. *Journal of Bacteriology*, 163, 88-93.

Fernandez, R. C., & Weiss, A. A. (1994). Cloning and Sequencing of a *Bordetella pertussis* Serum Resistance Locus. *Infection and Immunity*, 62(11), 4727-4738.

Fernandez, R. C., & Weiss, A. A. (1996). Susceptibilities of *Bordetella pertussis* strains to Antimicrobial Peptides. *Antimicrobial Agents and Chemotherapy*, 40(4), 1041-1043.

Ferry, N. S. (1910). A preliminary report of the bacterial findings in canine distemper. *Am Vet Rev*, 37(499-504).

Filion, R., Clautier, S., Vrancken, E. R., & Bernier, G. (1967). Respiratory infection in the turkey caused by a bacterium related to *Bordetella bronchiseptica. Canadian Journal of Comparative Medicine and Veterinary Science*, 31(129-134).

Fine, P. E. M. (1988). Epidemiological considerations for whooping cough eradication. In A. C. Wardlaw & R. Parton (Eds.), *Pathogenesis and Immunity in Pertussis* (pp. 451-467). John Wiley and Sons.

Fine, P. E. M., & Clarkson, J. A. (1987). Reflections on the efficacy of pertussis vaccines. *Reviews of Infectious Diseases*, 9, 866-883.

Finlay, B. B., & Falkow, S. (1997). Common themes in microbial pathogenicity revisited. *Microbiology and Molecular Biology Reviews*, 61(2), 136-169.

Finn, T. M., & Amsbaugh, D. F. **(1998)**. Vag8, a *Bordetella pertussis bvg*-regulated protein. *Infection and Immunity*, **66**, 3985-3989.

Finn, T. M., Shahin, R., & Mekalanos, J. J. **(1991)**. Characterisation of *vir*-activated tn*phoa* gene fusions in *Bordetella pertussis*. *Infection and Immunity*, **59**(9), 3273-3279.

Finn, T. M., & Stevens, L. A. **(1995)**. Tracheal colonisation factor: a *Bordetella pertussis* secreted virulence determinant. *Molecular Microbiology*, **16**(4), 625-634.

Forde, C. B., Parton, R., & Coote, J. C. **(1998)**. Bioluminescence as a reporter of intracellular survival of *Bordetella bronchiseptica* in murine phagocytes. *Infection and Immunity*, **66**(7), 3198-3207.

Friedman, R. L., Nordensson, K., Wilson, L., Akporiaye, E. T., & Yocum, D. E. **(1992)**. Uptake and intracellular survival of *Bordetella pertussis* in human macrophages. *Infection and Immunity*, **60**(11), 4578-4585.

Genetics Computer Group **(1991)**. Program manual for the GCG Package. *University of Wisconsin Genetics Computer Group, Madison*. Version 7

Goldberg, M. B., Barzu, O., Parsot, C., & Sansonetti, P. J. **(1993)**. Unipolar localisation and atpase activity of Icsa, a *Shigella flexneri* protein involved in intracellular movement. *Journal of Bacteriology*, **175**(8), 2189-2196.

Goldman, W. E. **(1988)**. Tracheal cytotoxin of *Bordetella pertussis*. In A. C. Wardlaw & R. Parton (Eds.), *Pathogenesis and Immunity in Pertussis* (pp. 231-246). John Wiley and Sons Ltd.

Goldman, W. E., & Herwaldt, L. A. **(1985)**. *Bordetella pertussis* tracheal cytotoxin. *Developments in Biological Standardization*, **61**, 103-111.

Grenfell, B. T., & Anderson, R. M. **(1989)**. Pertussis in England and Wales: an investigation of transmission dynamics and control by mass vaccination. *Proceedings of the Royal Society of London Series-B Biological Sciences*, **236**, 213-252.

Guiso, N., Rocancourt, M., Szatanik, M., & Alonso, J. M. **(1989)**. *Bordetella* adenylate cyclase is a virulence associated factor and an immunoprotective antigen. *Microbial Pathogenesis*, **7**, 373-380.

Hall, E., Parton, R., & Wardlaw, A. C. **(1994)**. Cough production, leucocytosis and serology of rats infected intrabronchially with *Bordetella pertussis*. *Journal of Medical Microbiology*, **40**, 204-213.

Halperin, S. A. **(1999)**. Developing better paediatric vaccines. *BioDrugs*, **12**(3), 175-191.

Hamstra, H. J., Kuipers, B., Schijf-Evers, D., Loggen, H. G., & Poolman, J. T. **(1995)**. The purification and protective capacity of *Bordetella pertussis* outer membrane proteins. *Vaccine*, **13**(8), 747-752.

Hamstra, H. J., & Poolman, J. T. (Unpublished - poster presentation). Development of a *Bordetella pertussis* OMP(s) based vaccine. Unit for Bacterial vaccine Development and Pathogenesis Research.

Hancock, R. E. W. **(1991)**. Bacterial outer membranes: evolving concepts. *ASM News*, **57**(4), 175-182.

Hanke, C., Hess, J., Schumacher, G., & Goebel, W. **(1992)**. Processing by OmpT of fusion proteins carrying the HlyA transport signal during secretion by the *Escherichia coli* transport system. *Molecular and General Genetics*, **233**, 42-48.

Hannah, J. H., Li, Z. M., Kessel, M., Steven, A. C., Nguyen, N. Y., Ewell, J. B., Manclark, C. R., & Brennan, M. J. **(1990)**. Characterization of the major outer membrane porin protein of *Bordetella pertussis*. In C. R. Manclark (Ed.), Proceedings of the Sixth International Symposium on Pertussis, (pp. 141-147). DHSS Publication Number (FDA) 90-164.

Hazenbos, W. L., van den Berg, B. M., Geuijen, C. W., Mooi, F. R., & van Furth, R. **(1995)**. Binding of FimD on *Bordetella pertussis* to very late antigen-5 on monocytes activates complement receptor type 3 via protein tyrosine kinases. *Journal of Immunology*, **155**(8), 3972-3978.

Hefferman, E. J., Harwood, J., Fierer, J., & Guiney, D. **(1992a)**. The *Salmonella typhimurium* virulence plasmid complement resistance gene *rck* is homologous to a family of virulence related outer membrane protein genes, including *pagC* and *ail*. *Journal of Bacteriology*, **174**(1), 84-91.

Hefferman, E. J. S., Kimura, J. H., & Hascall, V. C. **(1992b)**. Mechanism of resistance to complement-mediated killing of bacteria encoded by the *Salmonella typhimurium* virulence plasmid gene *rck*. *Journal of Clinical Investigation*, **90**, 953-964.

Heiss, L. N., Flak, T. A., Lancaster, J. R., McDaniel, M. L., & W.E., G. **(1994)**. Interleukin-1 is linked to the respiratory epithelial cytopathology of pertussis. *Infection and Immunity*, **61**, 3123-3128.

Henderson, I. R., Navarro-Garcia, F., & Nataro, J. P. **(1998)**. The great escape: structure and function of the autotransporter proteins. *Trends in Microbiology*, **6**(9), 370-378.

Hendrixson, D. R., de la Morena, M. L., Stathopoulos, C., & St Geme III, J. W. **(1997)**. Structural determinants of processing and secretion of the *Haemophilus influenzae* Hap protein. *Molecular Microbiology*, **26**(3), 505-518.

Hewlett, E. L., & Gordon, V. M. **(1988)**. Adenylate cyclase toxin of *B. pertussis*. In A. C. Wardlaw & R. Parton (Eds.), *Pathogenesis and Immunity in Pertussis* (pp. 193-209). John Wiley and Sons Ltd.

Hirst, T. R., & Welch, R. A. **(1988)**. Mechanisms for secretion of extracellular proteins by Gram-negative bacteria. *Trends in Biological Science*, **13**, 263-269.

Holland, B. **(1998)**. Autotransporters: protein contortionists whose carboxy termini translocate their own amino terminal domains. *Trends in Microbiology*, **6**(10), 388-389.

Horiguchi, Y., Inove, N., Masuda, M., Kashimoto, Y., Katahiri, J., Sugimoto, N. & Matsuda, M. **(1997)**. *Bordetella bronchiseptica* dermonecrotizing toxin induces reorganisation of actin stress fibres through deamination of Gln-63 of the GTP-binding protein Rho. *Journal of Biological Chemistry*, **94**, 11623-11626.

Hueck, C. J. **(1998)**. Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiology and Molecular Biology Reviews*, **62**(2), 379-433.

Hynes, R. O. **(1987)**. Integrins: a family of cell surface receptors. *Cell*, **48**, 549-554.

Iida, T., & Okonogi, T. **(1971)**. Lienotoxicity of *Bordetella pertussis* in mice. *Journal of Medical Microbiology*, **4**, 51-61.

Inouye, M., & Yee, M. **(1973)**. Homogeneity of envelope proteins of *Escherichia coli* seperated by gel electrophoresis in sodium dodecyl sulfate. *Journal of Bacteriology*, **113**(1), 304-312.

Irons, L. I., Ashworth, L. A. E., & Wilton-Smith, P. **(1983)**. Heterogeneity of the filamentous haemagglutinin of *Bordetella pertussis* studied with monoclonal antibodies. *Journal of General Microbiology*, **129**, 2769-2778.

Ishibashi, Y., Claus, S., & Relman, D. A. **(1994)**. *Bordetella pertussis* filamentous hemagglutinin interacts with a leukocyte signal transduction complex and stimulates bacterial adherence to monocyte CR3 (CD11b/CD18). *Journal of Experimental Medicine*, **180**, 1225-1233.

Jacob-Dubuisson, F., Buisine, C., Mielcarek, N., Clement, E., Menozzi, F. D., & Locht, C. **(1996)**. Amino-terminal maturation of the *Bordetella pertussis* filamentous haemagglutinin. *Molecular Microbiology*.

Jose, J., Jahnig, F., & Meyer, T. F. **(1995)**. Common structural features of IgA1 protease-like outer membrane protein autotransporters. *Molecular Microbiology*, **18**, 377-382.

Karimova, G., Bellalou, J., & Ullmann, A. **(1996)**. Phosphorylation-dependent binding of BvgA to the upstream region of the *cyaA* gene of *Bordetella pertussis*. *Molecular Microbiology*, **20**(3), 489-496.

Karimova, G., & Ullmann, A. **(1997)**. Characterisation of DNA binding sites for the BvgA protein of *Bordetella pertussis*. *Journal of Bacteriology*, **179**(11), 3790-3792.

Kerr, J. R. **(1999)**. Type III (contact dependent) secretion in Gram-negative bacteria. *Reviews in Medical Microbiology*, **10**(3), 155-164.

Kerr, J. R., Rigg, G. P., Mathews, R. C., & Burnie, J. P. **(1999)**. The *Bpei* locus encodes type III secretion machinery in *B. pertussis*. *Microbial Pathogenesis*, **27**, 349-367.

Kersters, K., Hinz, K. H., Hertle, A., Segers, P., A, L., Siegmann, O., & De Ley, J. **(1984)**. *Bordetella avium* sp. nov., isolated from the respiratory tracts of turkeys and other birds. *International Journal of Systematic Bacteriology*, **34**, 56-70.

Khelef, N., Danve, B., Quentin Millet, M. J., & Guiso, N. **(1993)**. *Bordetella pertussis* and *Bordetella parapertussis*: Two immunologically distinct species. *Infection and Immunity*, **61**(2), 486-490.

Kimura, M., & Kuno-Saki, H. **(1990)**. Developments in pertussis immuunisation in Japan. *Lancet*, **336**, 30-32.

Kinnear, S. M., Boucher, P. E., Stabitz, S., & Carbonetti, N. H. **(1999)**. Analysis of BvgA activation of the pertactin gene promoter in *B. pertussis*. *Journal of Bacteriology*, **181**(17), 5234-5241.

Klauser, T., Kramer, J., Otzelberger, K., Pohlner, J., & Meyer, T. F. **(1993)**. Characterisation of the *Neisseria* Igaß-core. the essential unit for outer membrane targeting and extracellular protein secretion. *Journal of Molecular Biology*, **234**, 579-593.

Klauser, T., Pohlner, J., & Meyer, T. F. **(1992)**. Selective extracellular release of cholera toxin B subunit by *Escherichia coli*: dissection of *Neisseria* Igaß-mediated outer membrane transport. *The EMBO Journal*, **11**(6), 2327-2335.

Kobisch, M., & Novotny, P. **(1990)**. Identification of a 68-kilodalton outer membrane protein as a major protective antigen of *Bordetella* bronchiseptica by using specific-pathogen-free piglets. *Infection anf Immunity*, **58**(2), 352-357.

Lacey, B. W. **(1960)**. Antigenic modulation in *Bordetella pertussis*. *Journal of Hygiene*, **58**, 57-93.

Lambert, H. J. **(1965)**. Epidemiology of a small pertussis outbreak in Kent County Michigan. *Public Health Report*, **80**(365-369).

Lambert-Buisine, C., Willery, E., Locht, C., & Jacob-Dubuisson, F. **(198)**. N-Terminal characterisation of the *Bordetella pertussis* filamentous haemogglutinin. *Molecular Microbiology*, **28**(6), 1283-1293.

Laemmli, U. K. **(1970)**. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685

Lapin, J. H. **(1943)**. Clinical manifestations. In C. C. Thomas (Ed.), *Whooping Cough* (pp. 3-7, 43-45, 112). Springfield.

Lazdunski, A. M. **(1989)**. Peptidases and proteases of *Escherichia coli* and *Salmonella typhimurium*. *FEMS Microbiology Reviews*, **63**, 265-276.

Lee, C. K., Roberts, A. L., Finn, T. M., & Mekalanos, J. J. **(1990)**. Invasion of HeLa 229, Chinese Hamster Ovary and U937 Cells by *Bordetella pertussis*. In C. R. Manclark (Ed.), Proceedings of the Sixth International Symposium on Pertussis, (pp. 115-125). DHSS Publication Number (FDA) 90-164.

Leigh, A. F., Coote, J. G., Parton, R., & Duggleby, C. J. (**1993**). Chromosomal DNA from both flagellate and non-flagellate *Bordetella* species contain sequences homologous to the *Salmonella* H1 flagellin gene. *FEMS Microbiology Letters*, **111**, 225-232.

Leininger, E., Ewanowich, C. A., Bhargava, A., Peppler, M. S., Kenimer, J. G., & Brennan, M. J. (**1992**). Comparative roles of the Arg-Gly-Asp sequence present in the *Bordetella pertussis* adhesins pertactin and filamentous haemagglutinin. *Infection and Immunity*, **60**(6), 2380-2385.

Leininger, E., Kenimer, J. G., & Brennan, M. J. (**1990**). Surface proteins of *Bordetella pertussis*: role in adhesion. In C. R. Manclark (Ed.), Proceedings of the Sixth International Symposium on Pertussis, (pp. 100-105). DHSS Publication Number (FDA) 90-164.

Leininger, E., Roberts, M., Kenimer, J. G., Charles, I. G., Fairweather, N., Novotny, P., & Brennan, M. J. (**1991**). Pertactin, an Arg-Gly-Asp-containing *Bordetella pertussis* surface protein that promotes adherence of mammalian cells. *Proceedings of the National Academy of Science USA*, **88**, 345-349.

Leyh, R., & Griffith, R. W. (**1992**). Characterisation of the outer membrane proteins of *Bordetella avium*. *Infection and Immunity*, **60**(3), 958-964.

Li, L. J., Dougan, G., Novotny, P., & Charles, I. G. (**1991a**). P.70 pertactin, an outer-membrane protein from *Bordetella parapertussis*: cloning, nucleotide sequence and surface expression in *Escherichia coli*. *Molecular Microbiology*, **5**(2), 409-417.

Li, Z. M., Hannah, J. H., Stibitz, S., Nguyen, N. Y., Manclark, C. R., & Brennan, M. J. (**1991b**). Cloning and sequencing of the structural gene for the porin protein of *Bordetella pertussis*. *Molecular Microbiology*, **5**(7), 1649-1656.

Lipscombe, M., Charles, I. G., Roberts, M., Dougan, G., Tite, J., & Fairweather, N. F. **(1991)**. Intranasal immunisation using the B subunit of the *Escherichia coli* heat-labile toxin fused to an epitope of the *Bordetella pertussis* P.69 antigen. *Molecular Microbiology*, **5**(6), 1385-1392.

Livey, L., Duggleby, C. J., & Robinson, A. **(1987)**. Cloning and nucleotide sequence analysis of the serotype 2 fimbrial subunit gene of *Bordetella pertussis*. *Molecular Microbiology*, **1**(2), 203-209.

Locht, C., Bertin, P., Menozzi, F. D., & Renauld, G. **(1993)**. The filamentous haemagglutinin, a mutifaceted adhesin produced by virulent *Bordetella* spp. *Molecular Microbiology*, **9**(4), 653-660.

Locht, C., & Cabezon, T. **(1990)**. Molecular biological studies on the structure-function relationship of pertussis toxin and filamentous hemagglutinin. In C. R. Manclark (Ed.), Proceedings of the Sixth International Symposium on Pertussis, (pp. 41-52). DHSS Publication Number (FDA) 90-164.

Lory, S. **(1992)**. Determinants of extracellular protein secretion in Gram-negative bacteria. *Journal of Bacteriology*, **174**(11), 3423-3428.

Loveless, B. J., & Saier, M. H. J. **(1997)**. A novel family of channel-forming, autotransporting, bacterial virulence factors. *Molecular Membrane Biology*, **14**, 113-123.

Magistris, M. T., Romano, M., Nuti, S., Rappuoli, R., & Tagliabue, A. **(1986)**. Dissecting human T-cell responses against *Bordetella* species. *Journal of Experimental Medicine*, **168**, 1351-1362.

Mahon, B. P., Ryan, M. S., Griffin, F., & Mills, K. H. G. **(1996)**. Interleukin-12 is produced by macrophages in response to live or killed *Bordetella pertussis* and enhances

the efficacy of an acellular pertussis vaccine by promoting induction of Th1 cells. *Infection and Immunity*, **64**(12), 5295-5301.

Makhov, A. M., Hannah, J. H., Brennan, M. J., Trus, B. L., Kocsis, E., Conway, J. F., Wingfield, P. T., Simon, M. N., & Steven, A. C. **(1994)**. Filamentous hemaglutinin of *Bordetella pertussis*. A bacterial adhesin formed as a 50-nm monomeric rigid rod based on a 19-residue repeat motif rich in beta strands and turns. *Journal of Molecular Biology*, **241**, 110-124.

Makoff, A. J., Oxer, M. D., Ballantine, S. P., Fairweather, N. F., & Charles, I. G. **(1990)**. Protective surface antigen P69 of *Bordetella pertussis*: Its characterisation and very high level expression in *Escherichia coli*. *Biotechnology*, **8**, 1030-1033.

Manetti, R., Massari, P., Burroni, D., de Bernard, M., Marchini, A., Olivieri, R., Papini, E., Montecucco, C., Rappuoli, R., & Telford, J. L. **(1995)**. *Helicobacter pylori* cytotoxin: importance of native conformation for induction of neutralising antibodies. *Infection and Immunity*, **63**, 4476-4480.

Mattoo, S., Miller, J. F., & Cotter, P. A. **(2000)**. Role of *Bordetella bronchiseptica* fimbriae in tracheal colonisation and developmant of humoral immune response. *Infection and Immunity*, **68**(4), 2024-2033.

Maurer, J., Jose, J., & Meyer, T. F. **(1997)**. Autodisplay: one-component system for efficient surface display and release of soluble recombinant proteins from *Escherichia coli*. *Journal of Bacteriology*, **179**(3), 794-804.

Maurer, J., Jose, J., & Meyer, T. F. **(1999)**. Characterization of the essential transport function of the AIDA-I autotransporter and evidence supporting structural predictions. *Journal of Bacteriology*, **181**(22), 7014-7020.

Marques, R.R., Carboneti, N.H. **(1997)**. Genetic analysis of pertussis toxin promoter activation in *Bordetella pertussis. Molecular Microbiology*. **26**(6). 1215-1224

Meerman, H. J., & Georgiou, G. **(1994)**. Construction and characterisation of a set of *E. coli* strains deficient in all known loci affecting the proteolytic stability of secreted recombinant proteins. *Biotechnology*, **12**, 1107-1110.

Melton, A. R., & Weiss, A. A. **(1989)**. Environmental regulation of virulence determinants in *Bordetella pertussis*. *Journal of Bacteriology*, **171**(11), 6206-6212.

Michiels, T., Wattiau, J. C., Brasseur, R., Ruysschaert, J. M., & Cornelis, G. R. **(1990)**. Secretion of Yop proteins by yersiniae. *Infection and Immunity*, **58**, 2840-2849.

Millis, L., Morris, C. A., Sheehan, M. C., Charlesworth, J. A., & Pussel, B. A. **(1993)**. Vitronectin-mediated inhibition of complement: evidence for different binding sites for C5b-7 and C9. *Clinical and Experimental Immunology*, **92**(114-119).

Miyazaki, H., Yanagidi, N., Horinouchi, S., & Beppu, T. **(1989)**. Characterisation of the Precursor of *Serratia marcescens* serine protease and COOH-terminal processing of the precursor during its excretion through the outer membrane of *Escherichia coli. Journal of Bacteriology*, **171**(12), 6566-6572.

Monji, N., Stebbins, M. R., McCoy, D. W., & Kuo, J. S.-C. **(1986)**. Isolation of the outer membrane components of *Bordetella pertussis* which enhance the immunogenicity of *Haemophilus influenzae* type b capsular polysaccharide polyrybosyl ribitol phosphate. *Infection and Immunity*, **51**(3), 865-871.

Montaraz, J. A., Novotny, P., & Ivanyi, J. **(1985)**. Identification of a 68-kilodalton protective antigen from *Bordetella bronchiseptica. Infection and Immunity*, **47**(3), 744-751.

Mooi, F. R., van Oirschot, H., Heuvelman, K., & van der Heide, H. G. J. **(1998)**. Polymorphism in the *Bordetella pertussis* virulence factors P.69/pertactin and pertussis toxin in the Netherlands. Temporal trends and evidence for vaccine-driven evolution. *Infection and Immunity*, **66**(2), 670-675.

Moran, A. P. **(1996)**. Bacterial surface structures - an update. *FEMS Immunology and Medical Microbiology*, **16**, 61-62.

Nakato, N., Tobe, T., Fukuda, L., Suzuki, T., Kokmatsu, K., Yoshikawa, M., & Sasakawa, C. **(1993)**. The absense of a surface protease, OmpT, determines the intercellular spreading ability of *Shigella*: the relationship between the ompT and kcpA loci. *Molecular Microbiology*, **9**(3), 459-468.

Navarro-Garcia, F., Sears, C., Eslava, C., Cravioto, A., & Nataro, J. P. **(1999)**. Cytoskeletal effects induced by Pet, the serine protease of enteroaggregative *Escherichia coli*. *Infection and Immunity*, **67**(5), 2184-2192.

Nikaido, H., & Saier, M. H. J. **(1992)**. Transport proteins in bacteria: Common themes in their design. *Science*, **258**, 936-942.

Novotny, P., Chubb, A. P., Cownley, K., & Montaraz, J. A. **(1985)**. Adenylate cyclase activity of a 68,000-molecular weight protein isolated from the outer membrane of *Bordetella bronchiseptica*. *Infection and Immunity*, **50**(1), 199-206.

O'Toole, P. W., Austin, J. W., & Trust, T. J. **(1994)**. Identification and molecular characterisation of a major ring-forming surface protein from the gastric pathogen *Helicobacter mustelae*. *Molecular Microbiology*, **11**(2), 349-361.

Pallen, M. J., & Wren, B. W. (**1997**). The HtrA family of serine proteases. *Molecular Microbiology*, **26**(2), 209-221.

Parton, R. (**1991**). Changing perspectives on pertussis and pertussis vaccination. *Reviews in Medical Microbiology*, **2**, 121-128.

Parton, R. (**1996**). New perspectives on *Bordetella* pathogenicity. *Journal of Medical Microbiology*, **44**, 233-235.

Parton, R. (**1998**). *Bordetella*. In Topley and Wilson (Ed.), *Microbiology and Microbial Infections*, **9**(2) 901-918. Edwars Arnold, London.

Parton, R., & Wardlaw, A. C. (**1975**). Cell-envelope proteins of *Bordetella pertussis*. *Journal of Medical Microbiology*, **8**, 47-57.

Patterson, M., & Handel, M. (Eds.). (**1998**). Trends guide to bioinformatics. Elsevier Science.

Perlman, D., & Halvorson, H. O. (**1983**). A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. *Journal of Molecular Biology*, **167**, 391-409.

Plotkin, S. A., & Cadoz, M. (**1997**). The acellular pertussis vaccine trials: an interpretation. *The Pediatric Infectious Disease Journal*, **16**(5), 508-517.

Pohlner, J., Halter, R., Beyreuther, K., & Meyer, T. F. (**1987**). Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease. *Nature*, **325**, 458-462.

Porter, J. F., Parton, R., & Wardlaw, A. C. **(1991)**. Growth and survival of *Bordetella bronchiseptica* in natural waters and in buffered saline without added nutrients. *Applied and Environmental Microbiology*, **57**(4), 1202-1206.

Porter, J. F., & Wardlaw, A. C. **(1993)**. Long-term survival of *Bordetella bronchiseptica* in lakewater and in buffered saline without added nutrients. *FEMS Microbiology Letters*, **110**, 33-36.

Poulain-Godefroy, O., Mielcarek, N., Ivanhoff, N., Remoue, F., Schacht, A., Phillips, N., Locht, C., Capron, A., & Riveau, G. **(1998)**. *Bordetella pertussis* filamentous hemagglutinin enhances the immunogenicity of liposome-delivered antigen administered intranasally. *Infection and Immunity*, **66**(4), 1764-1767.

Provence, D. L., & Curtiss III, R. **(1994)**. Isolation and characterisation of a gene involved in hemaglutination by an avian pathogenic *Escherichia coli* strain. *Infection and Immunity*, **62**(4), 1369-1380.

Pugsley, A. P. **(1993)**. The complete general secretory pathway in Gram negative bacteria. *Microbiological Reviews*, **57**(1), 50-108.

Pugsley, A. P., & Schwartz, M. **(1985)**. Export and secretion of proteins by bacteria. *FEMS Microbiology Reviews*, **32**, 3-38.

Rambow, A. A., Fernandez, R. C., & Weiss, A. A. **(1998)**. Characterisation of BrkA expression in *Bordetella bronchiseptica*. *Infection and Immunity*, **66**(8), 3978-3980.

Relman, D. A., Wright, S., D., Falkow, S., Saukkonen, K., & Tuomanen, E. **(1990)**. Mechanisms of *Bordetella pertussis* adherence to ciliated respiratory cells and human macrophages. In C. R. Manclark (Ed.), Proceedings of the Sixth International Symposium on Pertussis, (pp. 91-99). DHSS Publication Number (FDA) 90-164.

Renauld-Mongenie, G., Cornette, J., Mielcarek, N., Menozzi, F. D., & Locht, C. **(1996)**. Distinct roles of the N-Terminal and C-Terminal precursor domains in the biogenesis of the *Bordetella pertussis* filamentous hemagglutinin. *Journal of Bacteriology*, **178**(4), 1053-1060.

Ricci, S., Rappuoli, R., & Scarlato, V. **(1996)**. The pertussis toxin liberation genes of *Bordetella pertussis* are trancriptionally linked to the pertussis toxin operon. *Infection and Immunity*, **64**(4), 1458-1460.

Roberts, M., Fairweather, N. F., Leininger, E., Pickard, D., Hewlett, E. L., Robinson, A., Hayward, C., Dougan, G., & Charles, I. G. **(1991)**. Construction and characterisation of *Bordetella pertussis* mutants lacking the *vir*-regulated P.69 outer membrane protein. *Molecular Microbiology*, **5**(6), 1393-1404.

Roberts, M., Tite, J. P., Fairweather, N. F., Dougan, G., & Charles, I. G. **(1992)**. Recombinant P.69/pertactin: immunogenicity and protection of mice against *Bordetella pertussis* infection. *Vaccine*, **10**(1), 43-48.

Robinson, A., Irons, L. I., Seabrook, R. N., Pearce, A., Matheson, M., & Funnell, S. G. P. **(1990)**. Structure-function studies on *Bordetella pertussis* fimbriae. In C. R. Manclark (Ed.), Proceedings of the Sixth International Symposium on Pertussis, (pp. 126-135). DHSS Publication Number (FDA) 90-164.

Rostrand, K. S., & Esko, J. D. **(1997)**. Microbial adherence through peptidoglycans. *Infection and Immunity*, **65**(1), 1-8.

Ruoslachti, & E. Pierschbacher **(1986)**. Arg-Gly-Asp: A Versatile Cell Recognition Signal. *Cell*, **44**, 517-518

Salmond, G. P. C., & Reeves, P. J. **(1993)**. Membrane traffic wardens and protein secretion in Gram-negative bacteria. *Trends in Biological Science*, **18**, 7-12.

Sandros, J., & Tuomanen, E. **(1993)**. Attachment factors of *Bordetella pertussis*: mimicry of eukaryotic cell recognition molecules. *Trends in Microbiology*, **1**(5), 192-195.

Savelkoul, P. H. M., DeKerk, D. P. G., Willems, R. J., Mooi, F. R., Van-der-Zeijst, B. A. M., & Gaastra, W. **(1996)**. Characterisation of the *fim2* and *fim3* fimbrial subunit genes of *Bordetella bronchiseptica*: Roles of Fim2 and Fim3 fimbriae and flagella in adhesion. *Infection and Immunity*, **64**(12), 5098-5105.

Scarlato, V., Arico, B., Prugnola, A., & Rappuoli, R. **(1991)**. Sequential activation and environmental regulation of virulence genes in *Bordetella pertussis*. *The EMBO Journal*, **10**(12), 3971-3975.

Scarlato, V., Prugnola, A., Arico, B., & Rappuoli, R. **(1990)**. Positive transcriptional feedback at the *bvg* locus controls expression of virulence factors in *Bordetella pertussis*. *Proceedings of the National Academy of Science USA*, **87**, 6753-6757.

Schnaitman, C. A. **(1973)**. Outer membrane proteins of *Escherichia coli*. *Archives of Biochemistry and Biophysics*, **157**, 541 - 552.

See, Y. P., & Jackowski, G. **(1990)**. Estimating molecular weights of polypeptides by SDS gel electrophoresis. In T. E. Creighton (Ed.), *Protein Structure a practical approach* (pp. 1-19). Oxford University Press.

Shahin, R., Leef, M., Eldridge, J., Hudson, M., & Gilley, R. **(1995)**. Adjuvanticicty and protective immunity elicited by *Bordetella pertussis* antigens encapsulated in poly(DL-lactide-co-glycolide) microspheres. *infection and Immunity*, **63**(4), 1195-2000.

Shareck, F., & Cameron, J. **(1984)**. Cloning of *Bordetella pertussis* outer membrane proteins in *Escherichia coli*. *Journal of Bacteriology*, **159**(2), 780-782.

Shimada, K., Ohnishi, Y., Horinouchi, S., & Beppu, T. **(1994)**. Extracellular Transport of *Alcaligenes faecalis* in *Escherichia coli* Using the COOH-Terminal Domain of *Serratia macrcescens* Serine Protease. *Journal of Biochemistry*, **116**(2), 327-334.

Shine, J., & Dalgarno, L. **(1975)**. Determinant of cistron specificity in bacterial ribosomes. *Nature*, **254**, 34-38.

Stahl, S., & Uhlen, M. **(1997)**. Bacterial surface display: trends and progress. *Tibtech*, **15**, 185-192.

Stein, M., Kenny, B., Stein, M., & Finlay, B. B. **(1996)**. Characterisation of EspC, a 110-kilodalton protein secreted by Eenteropathogenic *E.coli* which is homologous to members of the immunoglobulin A protease-like family of secreted proteins. *Journal of Bacteriology*, **178**(22), 6546-6554.

Steinhauer, J., Agha, R., Pham, T., & Varga, A. W. **(1999)**. The unipolar *Shigella* surface protein IcsA is targetted directly to the bacterial pole: IcsP cleavage of IcsA occurs over the entire bacterial surface. *Molecular Microbiology*, **32**(2), 367-377.

St. Geme III, J. W., Morena, M. L. d. l., & Falkow, S. **(1994)**. A *Haemophilus influenzae* IgA protease-like protein promotes intimate interaction with human epithelial cells. *Molecular Microbiology*, **14**(2), 217-233.

Stibitz, S., Weiss, A. A., & Falkow, S. **(1988)**. Genetic analysis of a region of the *Bordetella pertussis* chromosome encoding filamentous hemagglutinin and the pleiotropic regulatory locus *vir*. *Journal of Bacteriology*, **170**(7), 2904-2913.

Strauch, K. L., Johnson, K., & Beckwith, J. (**1989**). Characterisation of *degP*, a gene required for proteolysis in the in the cell envelope and essential for growth of *Escherichia coli* at high temperature. *Journal of Bacteriology*, **171**(5), 2689-2696.

Strugnell, R., Dougan, G., Chatfield, S., Charles, I., Fairweather, N., Tite, J., Beesley, J., & Roberts, M. (**1992**). Characterisation of a *Salmonella typhimurium aro* vaccine strain expressing the P.69 antigen of *Bordetella pertussis*. *Infection and Immunity*, **60**(10), 3994-4002.

Stryer, L. (**1988**). Biochemistry (3rd ed.). W.H. Freeman and Company. New York

Stuyve, M., Moons, M., & Tommassen, J. (**1991**). Carboxy-terminal phenylalanine is essential for the correct assembly of a bacterial outer membrane protein. *Journal of Molecular Biology*, **218**, 141-148.

Suhr, M., Benz, I., & Schmidt, M. A. (**1996**). Processing of the AIDA-I precursor: removal of AIDA^{C} and evidence for the outer membrane anchoring as a β-barrel structure. *Molecular Microbiology*, **22**(1), 31-42.

Suzuki, T., Lett, M. C., & Sasakawa, C. (**1995**). Extracellular transport of VirG protein in *Shigella*. *Journal of Biological Chemistry*, **270**, 30874-30880

Towbin, H., Staehelin, T., & Gordon, J. (**1979**). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proceedings of the National Academy of Sciences, USA*. **76**, 4350-4354.

Tuomanen (**1988**). *Bordetella pertussis* adhesins. In A. C. Wardlaw & R. Parton (Eds.), *Pathogenesis and Immunity in Pertussis* (chapter 4). John Wiley and Sons.

Van Gijsegem, F., Genin, S., & Boucher, C. **(1993)**. Conservation of secretion pathways for pathogenicity determinants of plant and animal bacteria. *Trends in Microbiology*, **1**(5), 175-180.

Vandamme, P., Heyndrickx, M., Vancanneyt, M., Hoste, B., De Vos, P., Falsen, E., Kersters, K., & Hinz, K. H. **(1996)**. *Bordetella trematum* sp. nov., isolated from wounds and ear infections in humans, and reassessment of *Alcaligenes denitrificans* Ruger and Tan 1993. *International Journal of Systematic Bacteriogy*, **46**(4), 849-858.

Vandamme, P., Hommez, J., Vancanneyet, M., Monsieurs, M., Hoste, B., Cookson, B., Wirsing von Konig, C. H., Kersters, K., & Blackall, P. J. **(1995)**. *Bordetella hinzii* sp. nov., isolated from poultry and humans. *International Journal of Sytematic Bacteriology*, **45**, 37-45.

Veiga, E., de Lorenzo, V., & Fernandez, L. A. **(1999)**. Probing secretion and translocation of a β-autotransporter using a reporter single chain antibody Fv as a cognate passenger domain. *Molecular Microbiology*, **33**(6), 1232-1243.

Von Henge, G. **(1986)**. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Research*, **14**(11), 4683-4690.

Walker, K. E., & Weiss, A. A. **(1994)**. Characterisation of the dermonecrotic toxin in members of the genus *Bordetella*. *Infection and Immunity*, **63**(9), 3817-3828.

Wandersman, C. **(1992)**. Secretion across the bacterial outer membrane. *Trends in Genetics*, **8**(9), 317-322.

Wardlaw, A. C., Parton, R., & Hooker, M. J. **(1976)**. Loss of protective antigen, histamine-sensitising factor and envelope polypeptides in cultural variants of *Bordetella pertussis*. *Journal of Medical Microbiology*, **9**, 89-100.

Watanabe, M., Takimoto, H., Kumazawa, Y., & Amano, K. I. **(1990)**. Biological properties of lipopolysaccharides from *Bordetella pertussis*. *Journal of General Microbiology*, **136**, 489-493.

Weiss, A. **(1997)**. Mucosal immune defences and the response of *Bordetella pertussis*. *ASM News*, **63**(1), 22-28.

Weiss, A. A. **(1994)**. Chapter 22: Unorthodox secretion by Gram-negative bacteria. In V. L. Miller, J. B. Kaper, D. A. Portnoy, & R. R. Isberg (Eds.), *Molecular Genetics of Bacterial Pathogenesis* (pp. 341-349). Washington: American Society for Microbiology.

Weiss, M. S., Abele, U., Weckesser, J., Welte, W., Schiltz, E., & Schultz, G. E. **(1991)**. Molecular architecture and electrostatic properties of a bacterial porin. *Science*, **254**, 1627-1630.

Weiss, A. A., & Goodwin, M.S.M. **(1989)**. Lethal infection by *Bordetella pertussis* mutants in the infant mouse model. *Infection and Immunity*, **42**, 33-41.

Weyant, R. S., D.G., H., Weaver, R. G., Amin, M. F. P., Steigerwalt, A. G., O'Connor, S. P., Whitney, A. M., Daneshvar, M. I., & Brenner, D. J. **(1995)**. *Bordetella holmsii* sp. nov., a new gram-negative species associated with septicaemia. *Journal of Clinical Microbiology*, **33**, 1-7.

Willems, R. J. L., Kamerbeek, J., Geuijen, C. A. W., Top, J., Gielen, H., Gaastra, W., & Mooi, F. R. **(1998)**. The efficacy of a whole cell vaccine and fimbriae against *Bordetella pertussis* and *Bordetella parapertussis* in a respiratory mouse model. *Vaccine*, **16**(4), 410-416.

Willems, R. J. L., & Mooi, F. R. (**1996**). From whole cell to acellular pertussis vaccines. *Reviews in Medical Microbiology*, **7**(1), 13-21.

Wolff, J., Cook, G. H., Goldhammer, A. R., & Berkowitz, S. A. (**1980**). Calmodulin activates prokaryotic adenylate cyclase. *Proceedings of the National Acadamy of sciences of the USA*, **77**, 3841-3844.

Yuk, M. H., Harvill, E. T., & Miller, J. F. (**1999**). The BvgAS virulence control system regulates type III secretion in *Bordetella bronchiseptica*. *Molecular Microbiology*, **28**, 945-959.

Yuk, M. H., Harvill, E. T., Cotter, A., & Miller, J. F. (**2000**). Modulation of host immune responses, induction of apoptosis and inhibition of NF-kB activation by the *Bordetella* type III secretion system. *Molecular microbiology*, **35**(5), 991-1004.

6.0 Appendices

6.1 Appendix I

Media composition

Luria Bertani broth (LB)

1 litre

Tryptone	10g
Yeast extract	5g
Sodium chloride	10g

Add 1.2% agar for solid media

Cyclodextrin Liquid (CL) media

1 litre

Sodium-L-glutamate	10.7g
L-proline	0.24g
Sodium chloride	2.5g
Sodium di-hydrogen orthophosphate	0.5g
Magnesium chloride ($6H_2O$)	0.1g
Calcium chloride	0.02g
Tris	6.1g
Casamino acids	10g
Methyl-ß-clyclodextrin	0.25g
Potassium chloride	0.2g

pH to 7.6 and add vitamin solution to 0.5%

Vitamin Solution

75 ml

L-cysteine	0.04g
Iron sulphate ($7H_20$)	0.01g
Nicotinic acid	0.004g
Glutathione	0.15g
Ascorbic acid	0.4g

Casamino acid (CAA) solution

1 Litre

Casein hydrolysate	10g
Magnesium chloride ($6H_2O$)	0.1g
Calcium chloride	0.016g
Sodium chloride	5g

pH to 7.1; supplement with 20% glycerol for storage at -80°C

2x Yeast tryptone medium (2x YT)

1 Litre

Bacto-tryptone	16g
Bacto-yeast extract	10g
Sodium chloride	5g

SOC Medium

1 Litre

Tryptone	20g
Yeast extract	5g
Salts I (250 mM KCl, 1M NaCl)	10ml

After autoclaving add 10ml of 2M sterile glucose and 10ml of Salts II (1M $MgCl_2.6H_2O$, 1M $MgSO_4.7H_2O$)

Minimal Media (MM)

1x M9 salts (see below for 10x M9 salts recipe)
2% Casamino acids
0.2% Glucose
1mM $MgCl_2$

(Autoclave casamino acids then add M9 salts, glucose and $MgCl_2$)

10x M9 salts

1 Litre

Sodium dihydrogen orthophosphate	60g
Pottasium dihydrogen orthophosphate	30g
Sodium chloride	5g
Ammonium chloride	10g

pH 7.4

Molecular biology solutions

Step Solution

SDS	0.5% w/v
Tris	50mM (pH7.5)
EDTA	0.4M

Add 1 mg/ml proteinase K immediately before use.

Tris-Borate-EDTA (TBE) buffer. 5x stock solution

Tris	54g
Boric acid	27.5g
EDTA (0.5M stock)	20ml

DNA loading buffer. 6x stock solution

Tris	60mM
EDTA	6mM
Sucrose	40%
Bromophenol blue	0.25%

Southern blot solutions

Denaturing solution

Sodium hydroxide	0.5M

Sodium chloride	1M

Neutralising solution

Sodium chloride	1.5M
Tris	1M, pH 8.0

SSC. 20x stock solution

Sodium chloride	3M
Trisodium citrate	0.3M, pH 7.0

Chemiluminescence wash 1

dH_2O	890ml
x20 SSC	100ml
10%SDS	10ml

Chemiluminescence wash 2

dH_2O	895ml
x20 SSC	5ml
10%SDS	10ml

Buffer 1

Maleic acid	0.1M
Sodium chloride	0.15M

pH to 7.5

Buffer 2

Blocking stock solution (Boehringer Mannheim) (10%)	20ml
Buffer 1	180ml

Buffer 3

Tris-HCl	0.1M

Sodium chloride	0.1M
pH to 9.5	

Pre-hybridisation solution

Blocking stock solution (10%)	30ml
x20 SSC	75ml
Sarcosine (N-lauryl sarcosinate)	3ml
10% SDS	600µl
dH_2O	218.4ml

Washing buffer

Tween 20	0.3%
Buffer 1	500ml

Protein analysis solutions

Protein sample buffer

Glycerol	5ml
20% SDS	2.5ml
2-mercaptoethanol	0.5ml
Tris (0.5ml, pH 6.8)	2.5ml
Bromophenol blue	0.25%

Buffer A

Sodium phosphate	50mM, pH 7.4
Sucrose	5%

Stacking gel (6%)

Acrylamide/Bis solution	4ml
dH_2O	10.69ml
Tris-HCl (0.5M, pH 6.8)	5ml
20% SDS	100µl
TEMED	10µl
10% Ammonium persulfate (APS)	200µl

Separating gel (12%, adjust volume of Acrylamide/Bis solution and dH_2O \ to change percentage)

Acrylamide/Bis solution	4ml
dH_2O	12.6ml
Tris (1.5M, pH 8.8)	10ml
10% SDS	0.4ml
TEMED	30µl
10% APS	0.3ml

PAGE running buffer (10x)

1 Litre

Tris	30g
Glycine	144g
SDS	10g

pH 8.3

Coomassie gel stain

Coomassie blue	0.5g
Methanol	500ml
Acetic acid	100ml
dH_2O	400ml

For destain, omit the coomassie blue.

6.2 Appendix II

Sequence obtained from cosmid 3, which contained *bap-5*

```
   1 GATTGCGCGG CCGCGCCTGC TGCTGCTGGA CGAGCCGTCG ATGGGCCTGG CGCCGTTGAT
  61 GGTCGACAAG GTGTTCGAAG TGGTGCGCAC CATCGCGGCC GAGGGCGTGA CCATTCTGCT
 121 GATCGAGCAG AATGCCCGCC TGGCGCTCGA ACATGCCGAC CGCGGCTATG TGATGGAGTC
 181 GGGCGAGATC ACGCTGTCCG GCCCGGCCCG CGAGATGCTG CACGATCCCA AGGTGCGCGC
 241 GGCCTATCTG GGCGAAGTCG AGGCCNGAGC CAGTCCGCAC GCACGGGCGC CGGTTCTTCT
 301 TGCGGAACCG GCGCCCGTGC CGCGTTCAGG CTTCGGCCAG CGCCGTCGGG CGCGGCTTGC
 361 GATTGATCAT CAGGGCAATG ACCGCGGCGA TCAGGCCGGT GCTGCCGGCC AGTACGAAAG
 421 CCATCAGGTA GCTGCCCGTG GTTTCGCGTA CCACGCCGCC CATCCAGGCG GCGCTGGCCG
 481 CGCCCAACTG GTGGCCGGCG ACGATCCAGC CGAAGACGAT GGGGGCGTCG CGATCGCCAA
 541 AGGCTTCGGT GGTCAGGCGC AGTGTCGGCG GCACGGTGGC GATCCAGTCC AGCCCGAAGA
 601 AAATCGCGAA CAGCGACAGG CTGTAGAACG AGAAGTCCGA GTAGGGCAGG TACATCAGCG
 661 ACAGGCCGCG CAGGCTGTAA TAAACGAACA GCAGCCTGCG CGGGTCGTAG CGGTCGGTCA
 721 GCCAGCCCGA GGCCGTCGTG CCCACCAGGT CGAAGATGCC CATCAGCGCC AGCAGGCCGG
 781 CCGCCTGCAC TTCGGNNATG CCGNNGTNGC CGCACAGGGC GATCAGGTGC GTGNCCACCA
 841 GGCCATTGGT GGTAAAGCCG CACACGAAGA AGGTCGCGAA CAGATACCAG AACGTGCGGG
 901 TGCGCGCNGC GCNCCGCAAA GCGCCGAAGG TGGCGGCCAG CAGGCCGGTA CGCGGCGCGA
 961 CCGACGGCTC GGGCGCGTCC GGCGCGCTGC CATAGGAGCG CAGGCCGACA TCGGCGGGCC
1021 GGTCGGGCAC CAGCCACCAG GCAAGGGGCG CCATCAGCGC GGCGCCGGCA GCGACCGCCC
1081 AGACCACGCG GGTCCAGTCG CCGGAAGCCG CCAGGGCCGC CAGCACGGGC AGGAACAGCA
1141 GATTGCCGGT GGCCGTGCTG GNCGTCAGCA AGCCCATCAT CAGTCCGCGG CGCGTGGTGA
1201 ACCAGCGATT GACGATGGTC GCGCCCAGGA CCACCGCCAC CGCCCCCGAG CCTATCCCGG
1261 AGAACACGCC CCAGGTCAGC AGCAGGTGCC ACGACTCGGT CATGAAGGCG CTGGCCGCGC
1321 TGGAGGCGGC CATCAGCAGC AGGGCGCCGA TCAACACGCG GCGCAGGCCG AAGTGCTCCA
1381 TGGCGGCGGC GGCGAACGGG CCCGCCAGGC CATACANGAA AATACCGATC GCCGCNGCCA
1441 GNGAGATCGT GCTGCGGCTC CATCCGAAGG CCTGCTCCAG CGGCACGATC AGCACGCTGG
1501 GCGTGGAGCG CAGGCCGGCC GACACCAGCA GCGAGAAGAA GATCACGGCC ACCACGACGA
1561 AGGCGTACTT TTGTCCCAGG TGGCGAAACG GCGGTGGCAG GCGCGTTATA CTCATGTTAC
1621 TGACCGGTAC GTAGTGTGGA GCTCGCATCG TACGTACCGG TCAGTAACAT CGTCAAGCAG
1681 TTTTTGGCGA CCCCGATTTC CAGGAGCTCG CATGTCCCGT AAATCCGCTC CGTCATCCAC
1741 CGCCGCGCCG GCGCCCAAGC TGGCCGCCGA CCGTATCCGC GCCACCGCGC GCGAACTGTT
1801 CTATCGCGAA GGCATCCGCG CGGTGGGGGT CGACGCCATC GTATCGGCCG CCGGCGTCAC
1861 CAAGCCCAGC CTGTACCGCA GCTTCTCGTC CAAGGACGAA CTGGCCGCGT CGTACCTGCG
1921 CGANTACGAG ACCGAATTNT GGAGCCGCTT NGACGCCGGG GCGCAACGCC ATCCCGACGA
1981 TCCGCGCGCG CAGCTGCGGC TGTACCTGGA AAGCCTGGCC GAGCGCGCCG CGAACGCGGG
2041 GCCGTACCGC GGCTGCGGGC TGAGCAACGC CGCCGTGGAG TATCCCGATC CCGATCATCC
2101 GGCGCGTCTG GTGGCCGAAG CGCACAAGCG CGAACTGCGC CGTCGCCTGA TCGATATGGC
2161 GCGCCGCATG GGCGCGCGCG AGCCCGAAGT GCTCGGCGAC GGCTTGATGC TGCTCATCGA
2221 GGGCGCTTTC ATTTCGGGCC AGCTATTTCA CGGCGACGGG CCGGCGCGTC ACGTCGCGCG
2281 ATTGGCCGAC AGATTGATCG AAGCCAGTCT GTAGCGATTC CGTTTCATAT TCGTTCGATT
2341 TTCAAGCCGC CCCTCGGGGC GGTTTTTTTT GGTGGGGGAG GGCGATGTGC CGCGTGTAAG
2401 TCCGCCATTT TCGTACGTGT TCAGGTGCCG ATTGCCGGTA TCGGGATATA TGGGCTTATT
2461 TGAAACTAGA ACAGCTTTTT CACCGGGTGC CAGGCACCTG TTCGCAGCAG TTTTTTCACG
2521 ACTGTTTATT GGCCTCCTTT ACCGATTTGA ATATGAATGA CAGAAAATCC AATAGCATGT
2581 CGCCAGCGGC GAGATACTTG CGTTTCGGCG TTGTCGCGGT AGCGGGTGTG GCGGCGGGCG
2641 CGCTGCCTTC GTCGGATGTC GATGCCCAGG CCGCGCCGGC CGCCGCCGAG GTAGCCAAGA
2701 TCGAGGCTCT GTCGGATGCG GACATTTACA GCGACTACGA GCACGAGCAT GGCATCGTGA
2761 TGACGCCCGA TGGCAAGGAC GACTACATCA GTTACAGGTC CGCCGAGAGC GGTCGTCCGA
2821 AGCCCCCCCC CCCCCCTTTC AAACTTCAAA CCATTGGGTA ATGACGTCGT AGCGGAGCGG
2881 ATACGGGTAG AAGTGCACGG TGACGAAACC CTCGGCGTGT ATGTCGACTC GGAGCACCGT
2941 TCGCTCACCG TGCGTGACAG CACGATAGAT GCATACGGCA AGCCGCCTTC CGTCGACTCT
3001 CCCGATTATT ACGGTGCCGC GGCTGTCTAC GCCGGTACGC TGAATATCGA GAATTCCACG
3061 GTTCACCATA ACTATGCGGC CCAGCCGTTC GAAGACGCGG TAGGAGTCGG GGTAACCTCG
3121 CTCGGGGATA AGGCCATACT CAACGTTACC GACAGCGAGG TATCGGGTGC GAGGGGCGCG
3181 GTCATCGGTT GGGGGGGGGG GCGAAGCGAC ATTTACCGAT TCGGTCCTGC GTGGTTCGGC
3241 CTTCGGGCTG TACGCCGAAA TGTGCGACAC CTGCAGAGAT GATGATGGCA CCTCGCCTTC
3301 GATTCGCGTC CAAGGCGGGG TTGTTCAGGG CGGCATGGGT GCAAATAACG TCGCTGTGGT
3361 GGCAACAGGG TCTGGAAAGG TCGCGATCGA GAATGCGGAA CTGCTCGGAG CCAGCGGCAT
```

```
3421 GTACGCCACG TTCGGCGCGC AGGTCGATAT GAAAGGCGGG CGCATTCTGG CGCACAACAC
3481 CAATATCCTG GGAAGCCAGG GTTACGCCGA TGGTCCCTAT GGCGGCGTGG TCGTGACAGA
3541 GGACGGTCAA GTCAACCTGG AGGGCGCCAA GGTCAGTGCA ACTGGCCTGG GGGCCGCCGG
3601 CTTGTGGTTG CTGGGCGACA AGGACACCAG CCCGCGAGCC AGCCTGCGCA ACACCGACGT
3661 CCACGGAGAG GTCGCCGCCA TTGCGCTGGG GTTCAATGGC GAGGCGAACA TCTCGGGCGG
3721 CAGCTTGAGC GTAGAGGATG GGGCCGTGCT CACCACCCTG ACGCCCGATG CAGTCGAGTA
3781 TTACTACGAC TACGCCTTGT CCATGGAGCA TCTGCCAGCT GATGCGCCGT TGACGCCGGT
3841 CCGCGTCACG CTGTCCGATG GCGCGCGCGC CAGCGGAGAA ACGTTGATCG CGCATGGCGG
3901 GTTGTTGCCC ATGACGCTGC GCTTGAGCAG CGGGGTCGAC GCCCGCGGCG ACATCGTCAC
3961 GCTGCCGCCT TCCGCGCCGC CCGATTCCGC GGAGCAACCG GATGCCGAGC CGGAACCGGA
4021 TGCCGAGCTG GAACCGGACG CCGCGGCGCA GTCGGACGCC AAGGCGAATG CGCGGGTCAT
4081 GGCGCAGGTA GATGGCGGGG AACCTGTTGC CGTGCCGATC CCGGCCCCTT CGCATCCCGA
4141 TGCCCCGATC GACGTGTTCA TCGACAGCGG TGCCCAATGG CGGGGCATGA CCAAGACCGT
4201 CAATGCGTTG CGCATCGAGG ACGGCACCTG GACCGTCACC GGGTCGTCCA CGGTGAACAG
4261 CCTGCACCTG CAGGCAGGCA AGGTGGCGTA CGCAACGCCT GCCGAAAGCG ACGGAGAATT
4321 CAAACACCTG CGGGTCAAGA CCCTCTCGGG AAGCGGCCTG TTCGAGATGA ACGCCAGCGC
4381 CGACCTGAGC GATGGCGACC TGCTGGTCGT GTCCGACGAG GCCAGCGGGC AGCACAAGGT
4441 GCTGGTGCGA GGAGCCGGCA CGGAACCCAC CGGTGTGGAA AGCCTGACGC TGGTCGAGCT
4501 GCCCGAGGGC AGCCAGACGA AGTTCACGCT TGCCAACCGG GGCGGGGTGG TCGACGCCGG
4561 CGCGTTCCGC TATCGCCTGA CGCCGGACAA CGGTGTCTGG GGCCTGGAAC GGACCAGCCA
4621 GCTTTCGGCC GTCGCCAACG CGGCCTTGAA TACCGGGGGC GTGGGCGCGG CCAGCAGCAT
4681 CTGGTATGCG GAAGGCAATG CGCTCTCCAA GCGCCTGGGC GAGTTGCGGC TCGATCCCGG
4741 CGCGGGCGGC TTCTGGGGGC GCACGTTCGC CCAGAAGCAG CAGCTCGACA ACAAGGCTGG
4801 CCGACGCTTC GACCAGAAGG TGTACGGTTT CGAGCTGGGG GCCGACCATG CCATCGCAGG
4861 ACAGCAAGGG CGCTGGCACG TGGGCGGCCT GCTGGGCTAT ACCCGCGCAA GGCGCAGCTT
4921 CATCGATGAC GGCGCCGGGC ATACCGACAG CGCGCATATC GGGGCCTACG CGGCGTACGT
4981 GGCGGACAAC GGCTTCTATT TCGATTCGAC CCTGCGCGCC AGCCGCTTCG AGAACGACTT
5041 CACGGTAACG GCCACCGACG CCGTTTCCGT ACGGGGCAAG TACCGGGCCA ATGGGGTAGG
5101 CGCCACCTTG GAGGCCGGCA AACGTTTCAC GTTGCACGAC GGCTGGTTCG TCGAACCTCA
5161 GTCCGAGGTG TCGCTGTTCC ATGCCAGCGG CGGAACCTAC CGTGCCGCGA ACAACCTGTC
5221 GGTCAAGGAC GAAGGCGGCA CCTCCGCCGT GCTGCGCCTG GGCTTGGCGG CCGGGCGACG
5281 CATCGACCTG GGCAAGGACC GCGTGATCCA GCCCTATGCC ACCCTGAGCT GGCTGCAGGA
5341 ATTCAAAGGC GTCACGACCG TTCGCACCAA CGGGTACGGG CTGCGCACCG ACCTGAGCGG
5401 TGGCCGGGCT GAATTGGCGC TGGGCCTGGC CGCCGCGTTG GGGCGCGGCC ACCAGCTCTA
5461 CACTTCGTAC GAGTACGCCA AGGGCAACAA GCTGACCTTG CCTTGGACGT TCCACCTGGG
5521 CTATCGCTAC ACCTGGTAGC GGCGCATCGC AACAACGGGG TCATCGCGAC CCCGTTGTTG
5581 CATTCCCGGA CCG
```

6.3 Appendix III

Mouse lung counts after immunisation with various antigens from *B. pertussis* and intranasal challenge with either strains Taberman or 18-323.

Experiment	Challenge	Antigen preparation	Total lung counts from individual mice
R12	Taberman	ACV	<100, 500, <100, C*, C
		P.69	$6.7x10^3$, 150, $8.0x10^3$, 550, $1.0x10^5$
		Bap-5 urea	$3.0x10^5$, $3.4x10^5$, >$3.0x10^6$, $5.5x10^5$, $2.2x10^5$
		NTS	$9.6x10^5$, $1.1x10^6$, $2.8x10^6$, $1.5x10^6$, $1.9x10^6$
		E. coli lysate	>$3.0x10^6$, $2.5x10^6$, $1.7x10^6$, $1.4x10^6$, $3x10^5$
		Al	$7.7x10^5$, $5.3x10^6$, $1.4x10^6$, >$3x10^6$, C
	18-323	ACV	<100, 150, 200, C, C
		P.69	$2.8x10^3$, $5.0x10^4$, <100, C
		Bap-5 urea	$1.4x10^6$, $1.4x10^6$, $2.3x10^6$, $2.9x10^6$, $1.1x10^6$
		NTS	$1.3x10^6$, $1.5x10^6$, $1.5x10^6$, $1.5x10^6$, $1.7x10^6$, C
		E. coli lysate	$2.8x10^6$, $2.0x10^6$, >$3x10^6$, >$3x10^6$, >$3x10^6$,
		Al	$3.9x10^5$, $7.5x10^5$, $3.5x10^5$, >$3x10^6$, >$3x10^6$,
R13	Taberman	ACV	$1.2x10^5$, <100, <100, <100, $4.0x10^4$
		PCT	$1.8x10^6$, C, >$3x10^6$, >$3x10^6$, >$3x10^6$,
		BCT	$2.7x10^6$, >$3x10^6$, C, $2.8x10^6$, $9x10^5$

Appendix III (continued)

Experiment	Challenge	Antigen preparation	Total lung counts from individual mice
R14	Taberman	TCT	$4.4x10^5$, $2.9x10^6$, $7.3x10^5$, $3.4x10^5$, $1.9x10^6$
		Bap5CT	$2.1x10^6$, $1.5x10^6$, $1.6x10^6$, $5.1x10^6$, $2.4x10^6$
		Al	$>3x10^6$, $7.2x10^5$, $>3x10^6$, $1.2x10^6$, $>3x10^6$,
		ACV	<100, $1.0x10^5$, <100, $2.3x10^5$, $1.6x10^5$
		P.69	<100, $3.7x10^6$, $1.2x10^6$, $2.6x10^5$, $5x10^4$
		P.69(urea)	$1.2x10^5$, $1.0x10^4$, $2.0x10^4$, $4.4x10^5$, $1.2x10^6$
		Bap-5(PBS)	$9.0x10^6$, $1.0x10^7$, $5.0x10^6$, $2.7x10^6$, $1.0x10^6$
		Bap-5(urea)	$1.5x10^7$, $4.6x10^6$, $3.1x10^6$, $1.7x10^6$, $1.1x10^6$
		Bap-5(dialysed)	$5.5x10^6$, $5.2x10^6$, $2.3x10^6$, $2.7x10^6$, $2.3x10^6$
		Al	$4.2x10^6$, $6.0x10^6$, $2.0x10^5$, $4.0x10^5$, $3.5x10^6$
		GP	$2.6x10^7$, $1.6x10^7$, $1.0x10^6$, $3.4x10^6$, $1.3x10^6$

* C = Plate contaminated, no count possible.